Machine Learning Applications in MicroBooNE LArTPC Detector

Fermilab Machine Learning Group Meeting

Kazuhiro Terao @ Nevis, Columbia University
Taritree Wongjirad @ MIT

Machine Learning Applications
in
MicroBooNE
LArTPC Detector

Fermilab Machine Learning Group Meeting

Outline

- Machine learning apps in MicroBooNE
- LArTPC image data + challenges
- Convolutional Neural Networks in MicroBooNE
- Summary

Machine Learning Applications in UB

Boosted Decision Tree

- Used for low energy (>40 MeV) NC1P search
- Input: reconstructed parameters (length, angle, etc...)
- Analysis details available in UB public note page

Machine Learning Applications in UB

Boosted Decision Tree

- Used for low energy (>40 MeV) NC1P search
- Input: reconstructed parameters (length, angle, etc…)
- Analysis details available in UB public note page

Convolutional Neural Networks

- Demonstrated with LArTPC in 1st UB publication
- Usage being developed for multiple purpose
  - Reconstruction: vertex detection, PID, clustering…
  - Analysis: final state classifier
- Input: either raw data (waveforms) or reconstruction

Machine Learning Applications in UB

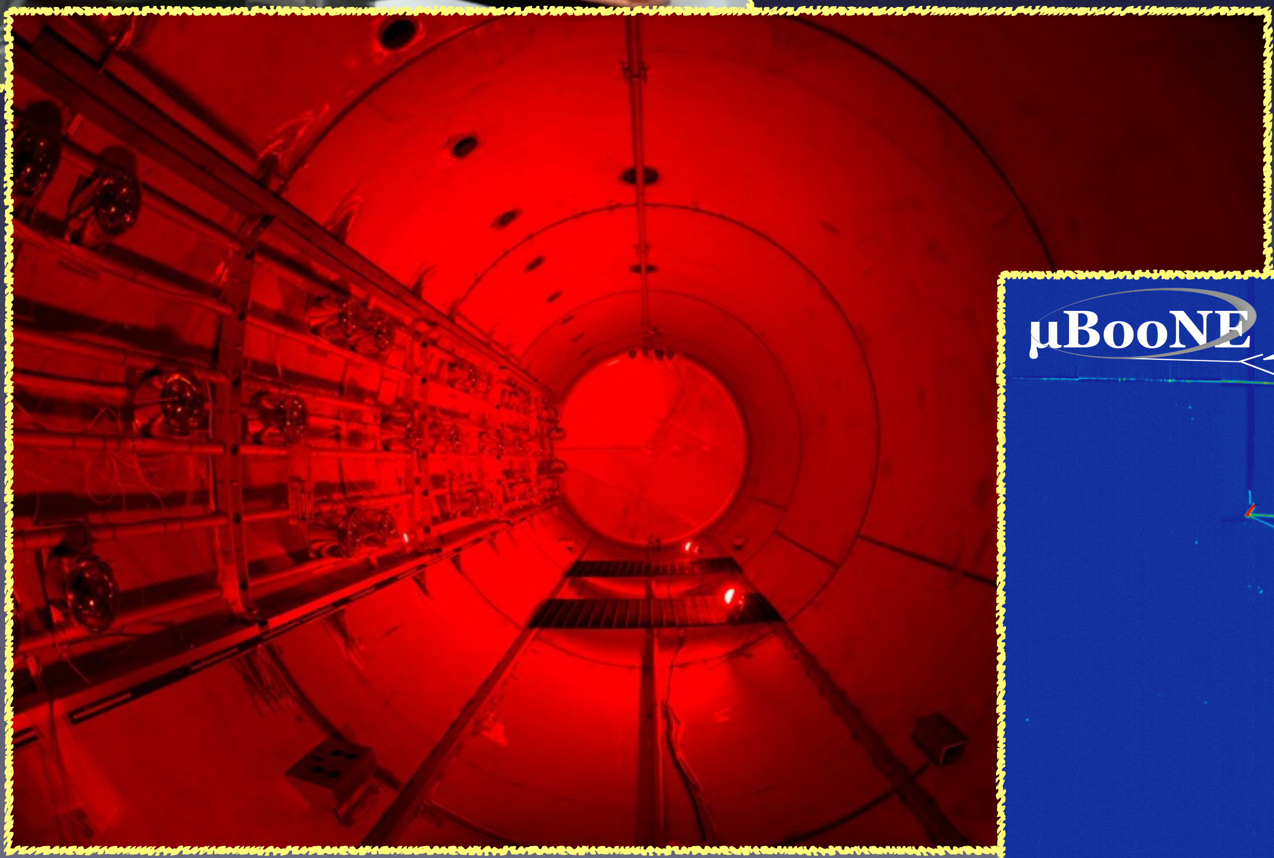
Boosted Decision Tree

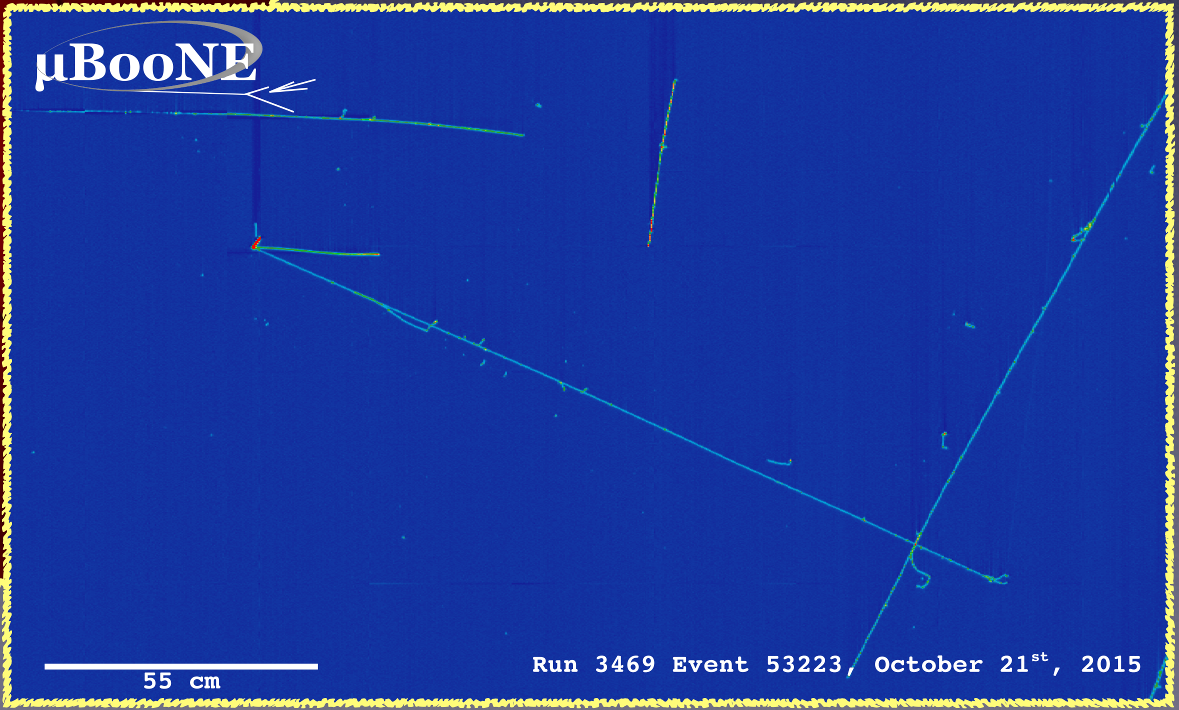
- Used for low energy (>40 MeV) NC1P search
- Input: reconstructed parameters (length, angle, etc...)
- Analysis details available in UB public note page

Convolutional Neural Networks

- Demonstrated with LArTPC in 1st UB publication
- Usage being developed for multiple purpose
  - Reconstruction: vertex detection, PID, clustering...
  - Analysis: final state classifier
- Input: either raw data (waveforms) or reconstruction

My focus today





MicroBooNE
LArTPC
Image Data

MicroBooNE!

MicroBooNE

Physics goal: **understand excess ν_e observed by MiniBooNE**

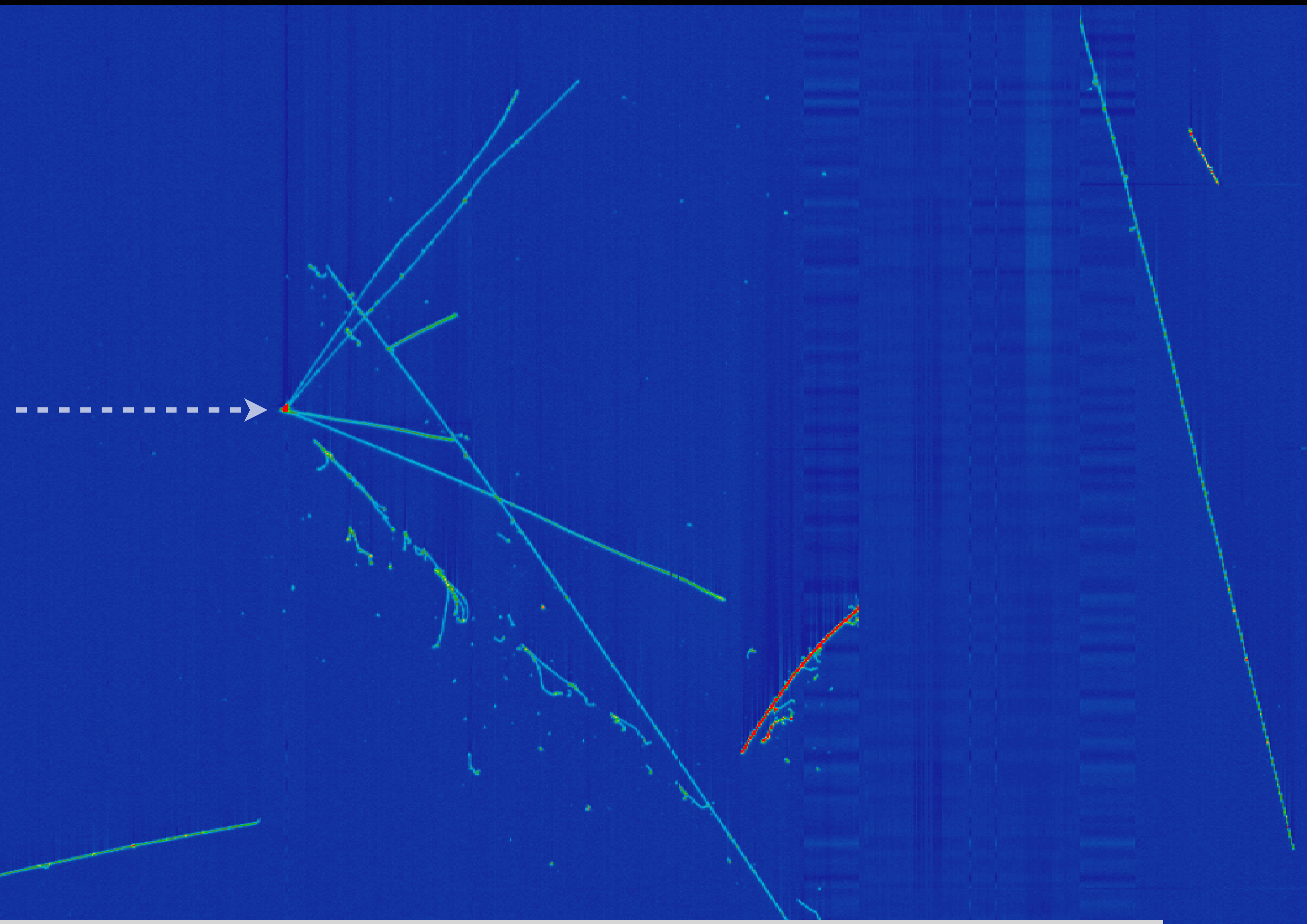
- Must be able to identify ν_e events at low energy (100 to 600 MeV)

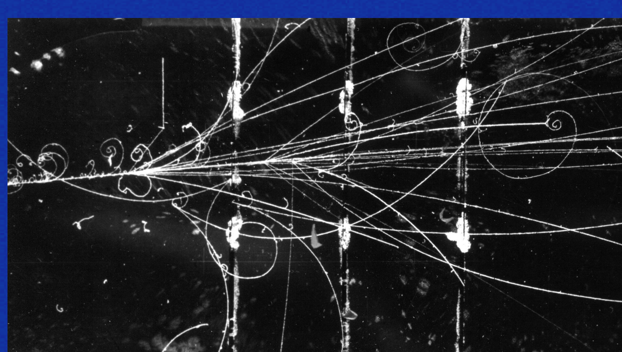
… **and more**:

- LArTPC R&D, event reconstruction, ν-Ar x-section & nuclear effects

LArTPC: Particle Imaging Machine

µBooNE



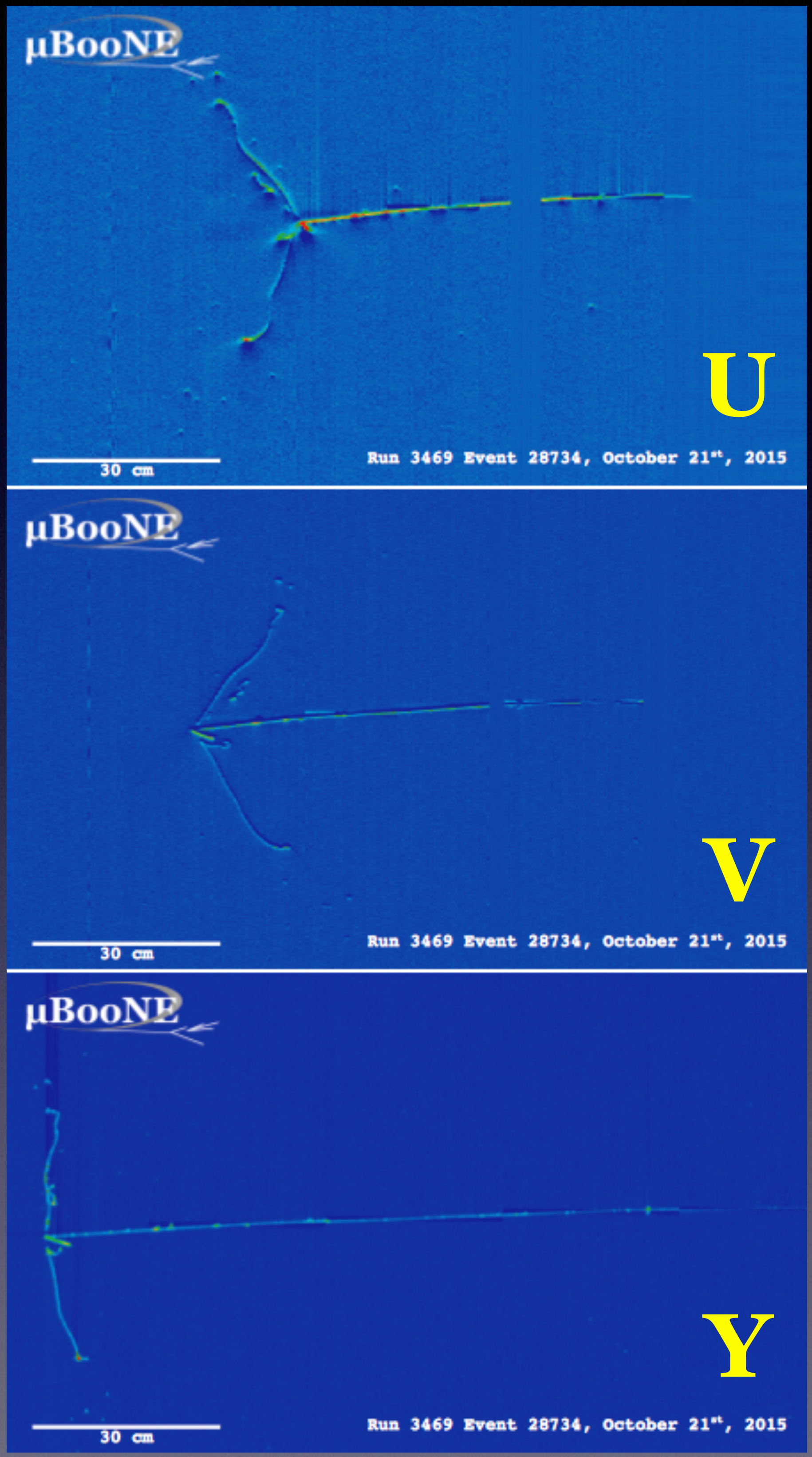


Bubble Chamber

Liquid Argon Time Projection Chamber

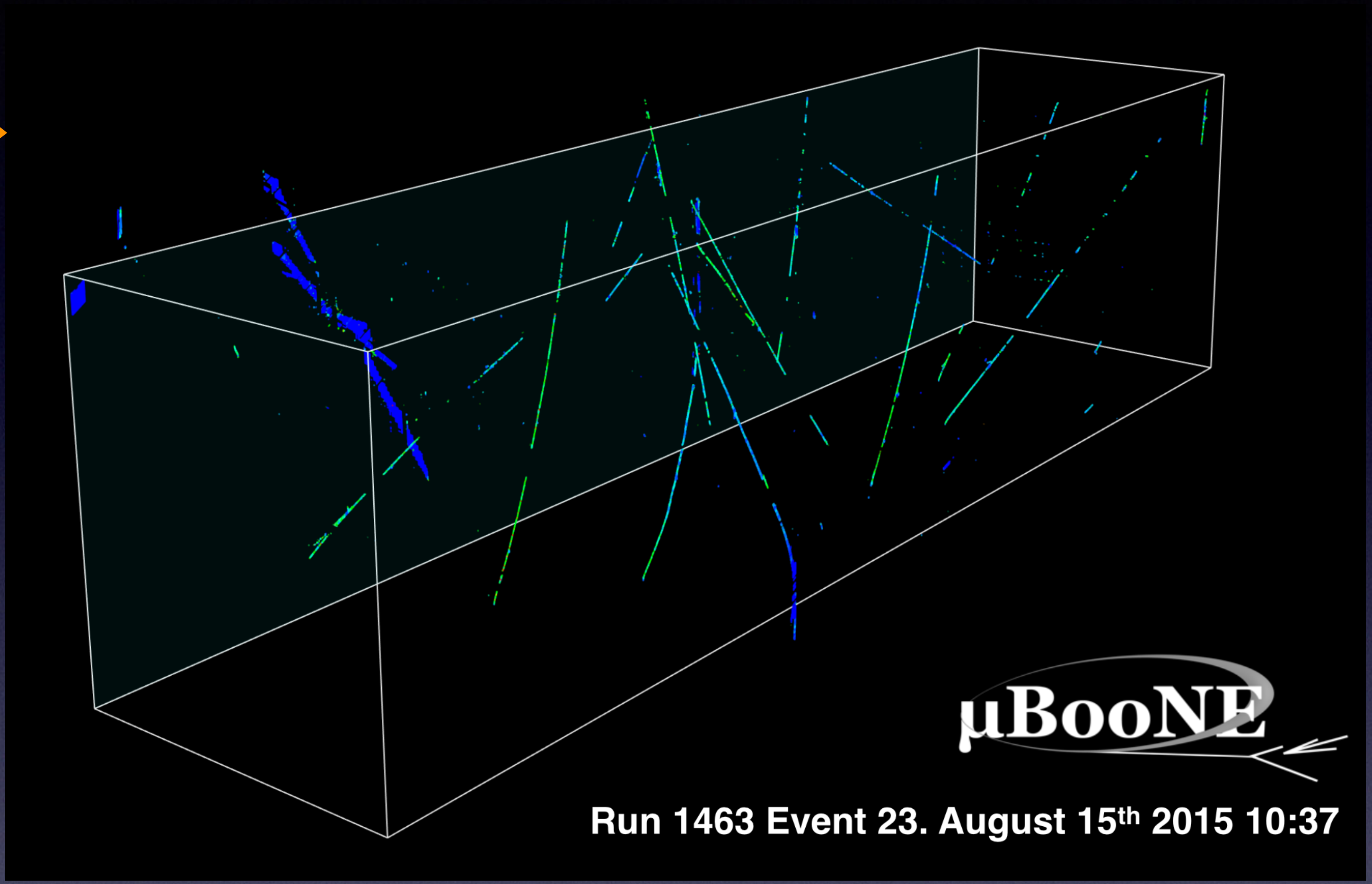
- Digitized bubble Chamber-like images
- Hi-resolution (~3 mm/px) 2D views + calorimetric information

LArTPC: Particle Imaging Machine



Three 2D Views

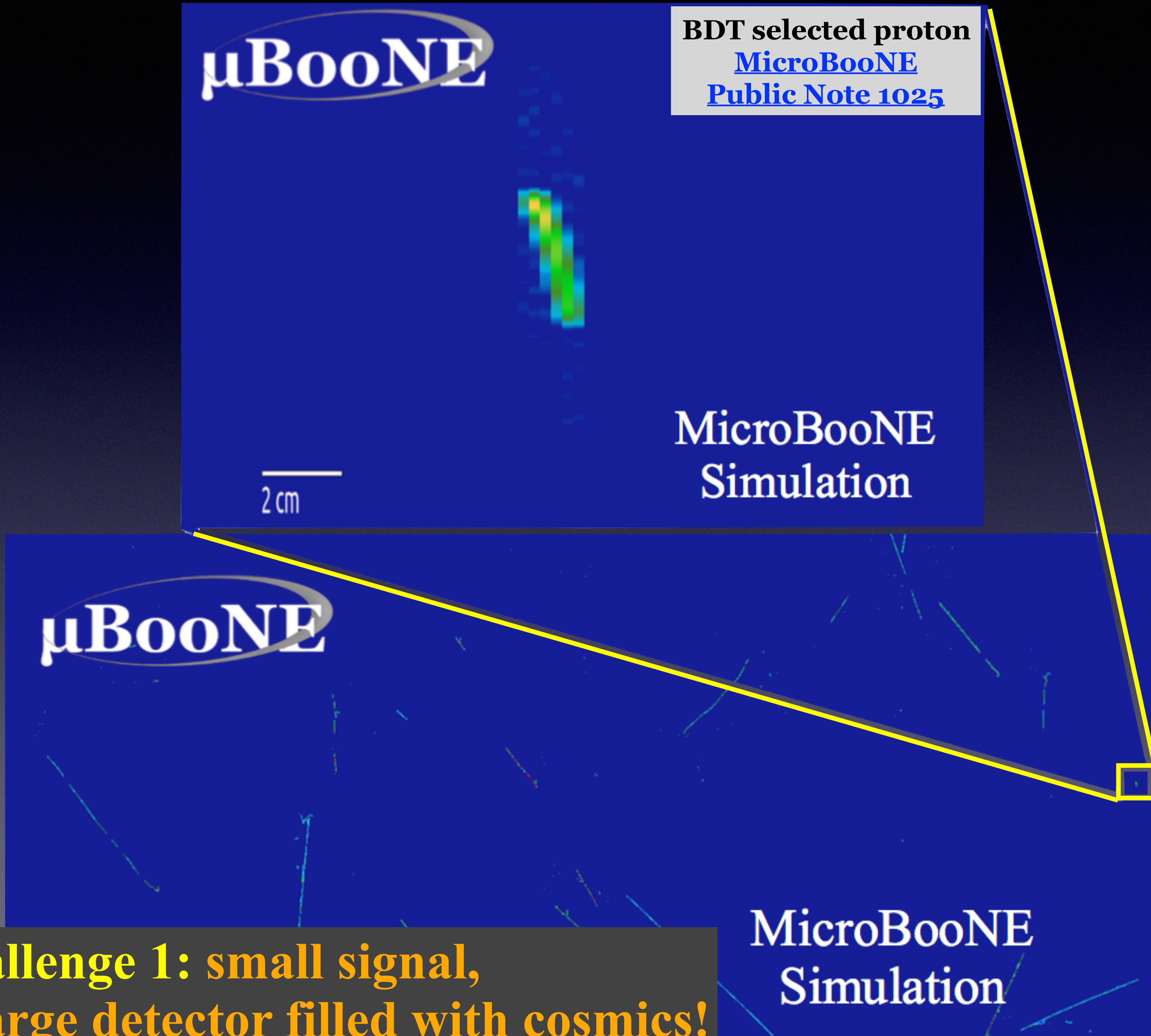
Reconstructed 3D View



Challenges

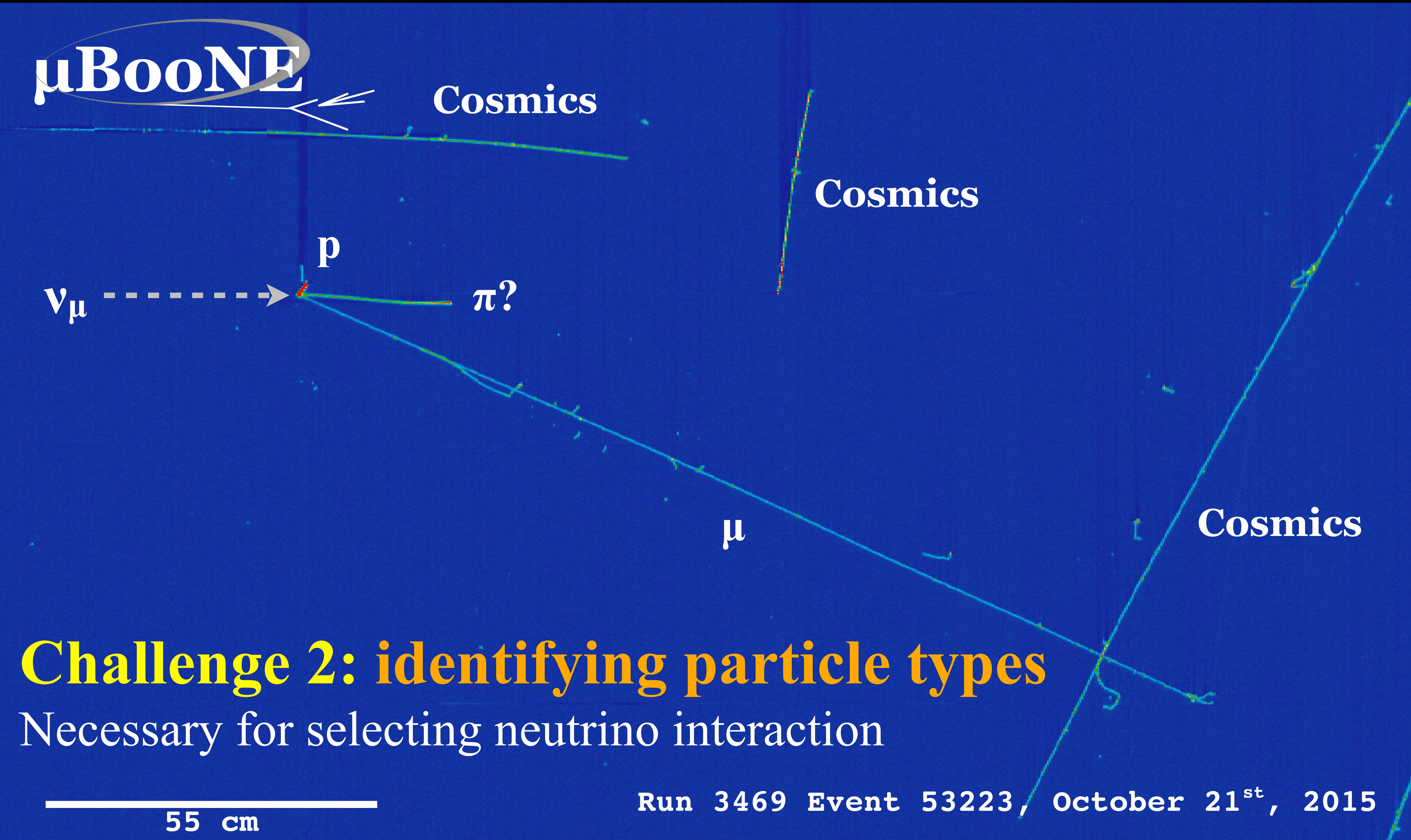
- Complicated event reconstruction
- Small signal, large detector, high rate of un-tagged cosmics

Challenges for Neutrino Analysis (I)



Challenge 1: small signal,
a large detector filled with cosmics!

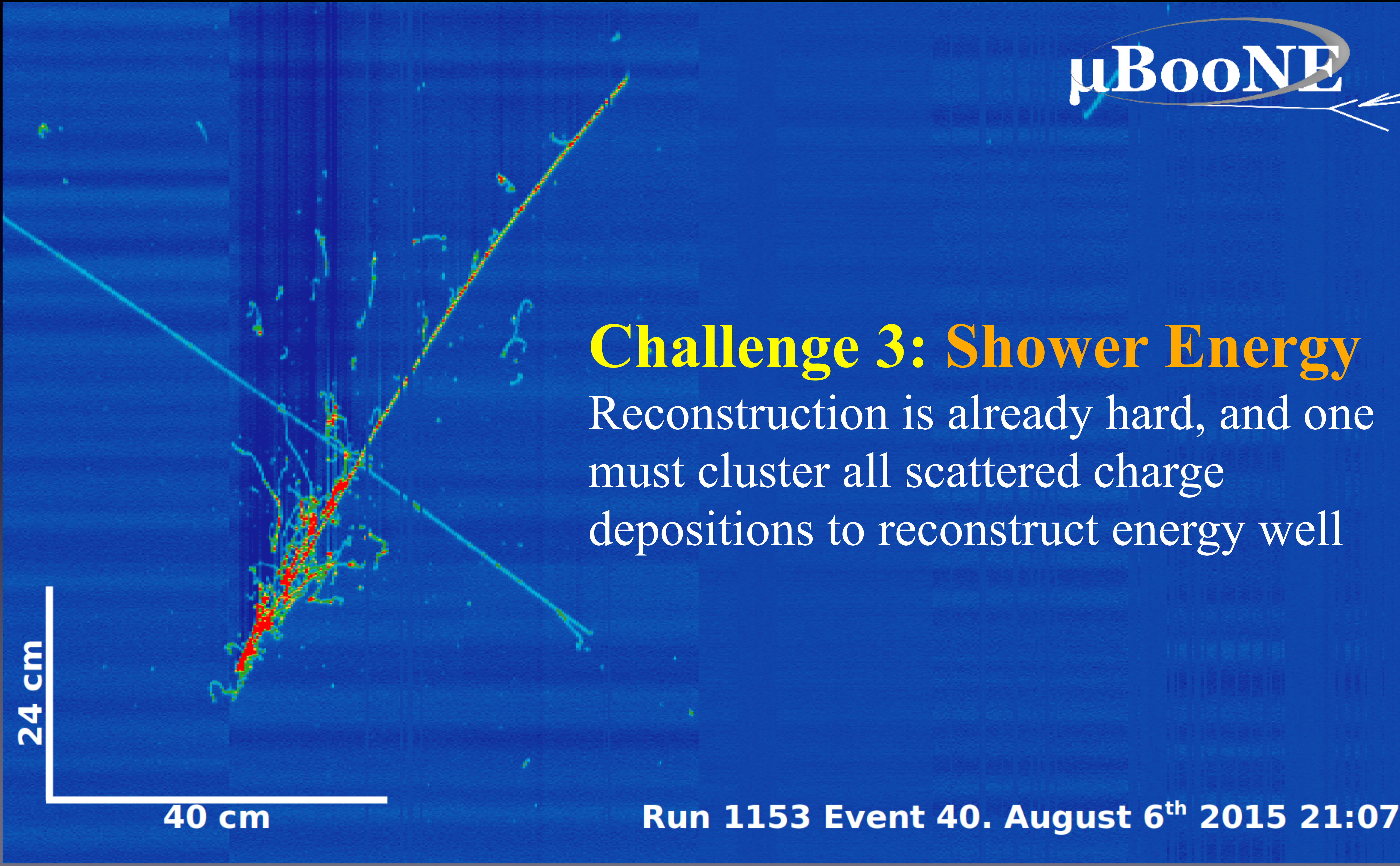
Challenges for Neutrino Analysis (II)



Challenge 2: identifying particle types

Necessary for selecting neutrino interaction

Challenges for Neutrino Analysis (III)



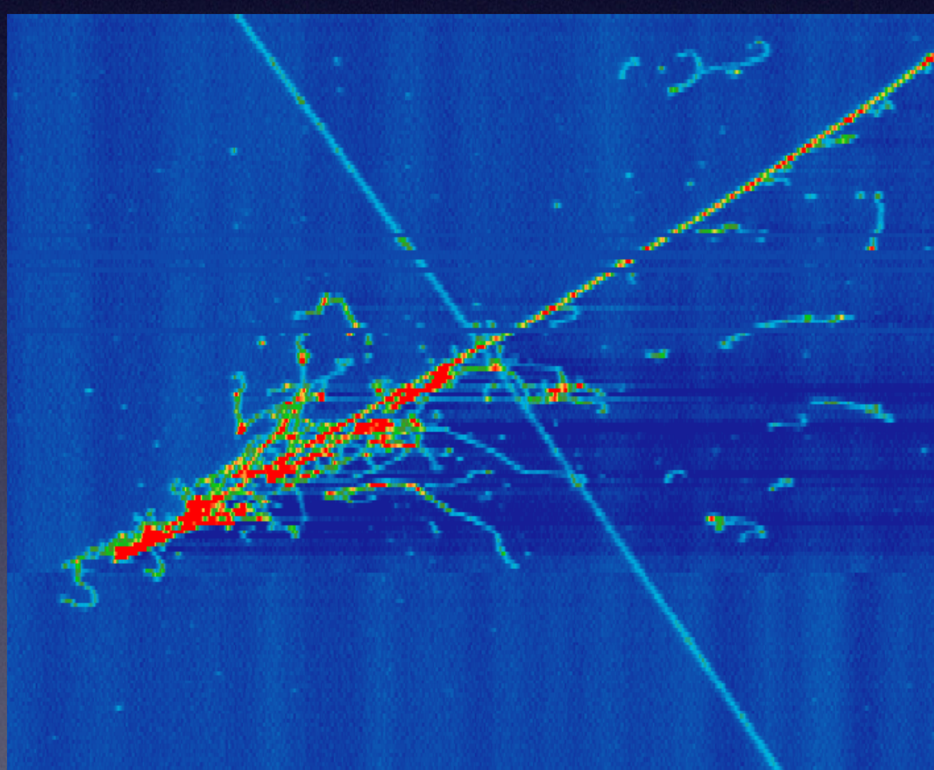
Challenge 3: Shower Energy

Reconstruction is already hard, and one must cluster all scattered charge depositions to reconstruct energy well

Challenges for Neutrino Analysis (IV)

Challenge 4: programming is not easy

Need efficient, fast pattern recognition algorithms and a framework to run a chain (or multiple chains) of them



Our data is an "image", a matrix of numbers

we wish

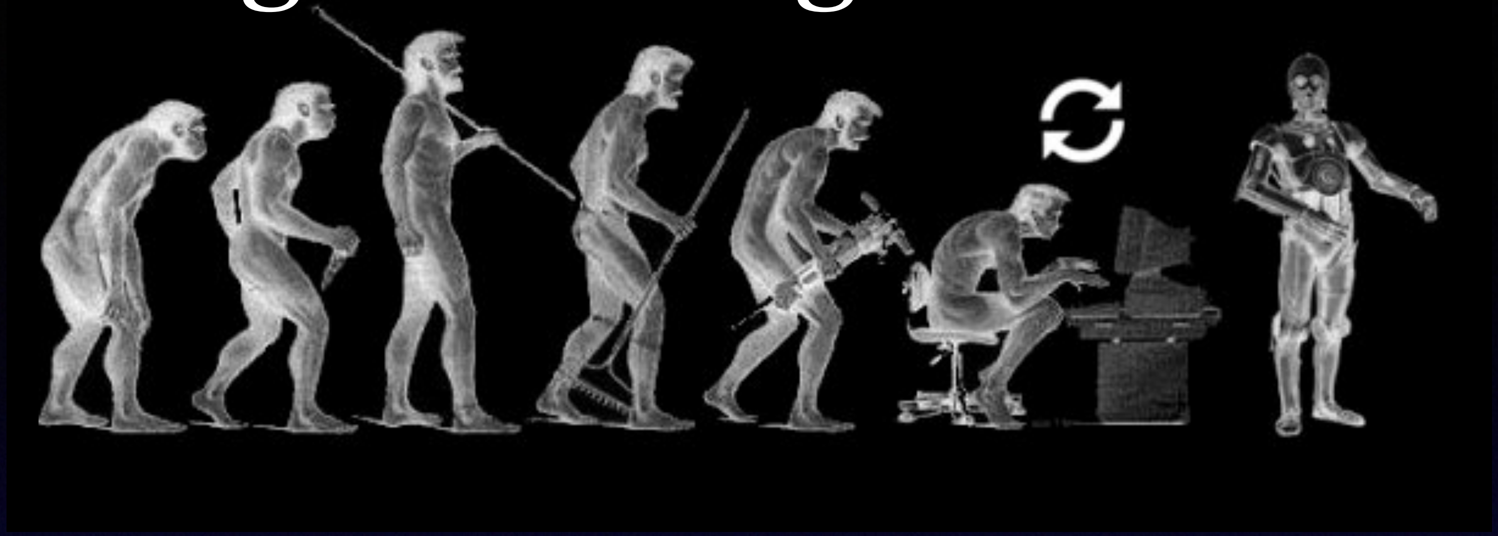
Not how it looks in C++

in reality

01101010100101011010101001011010
10111010101001010100010010101101
0101001011010101001010110101010
01011010101001010110101010101101
0101001010110101010010110101010
01011010101001010110101010010110
10101001010110101010101101010100
10101101010100110101101010100101

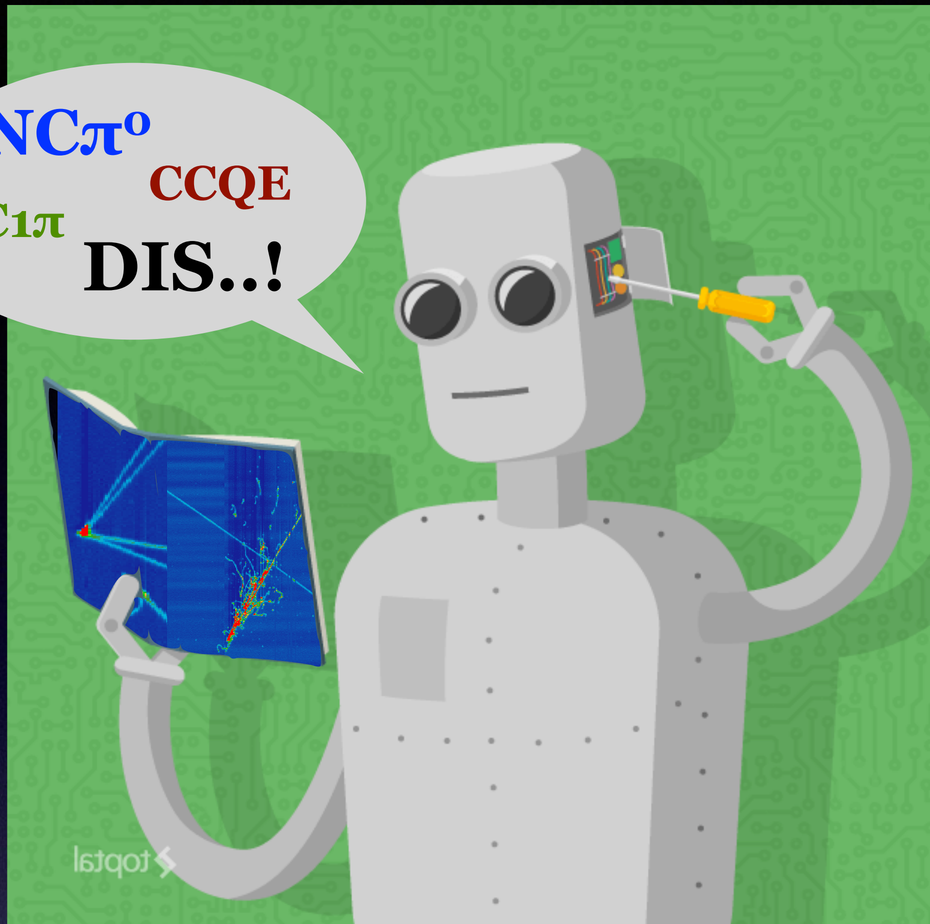
This is how it looks in C++

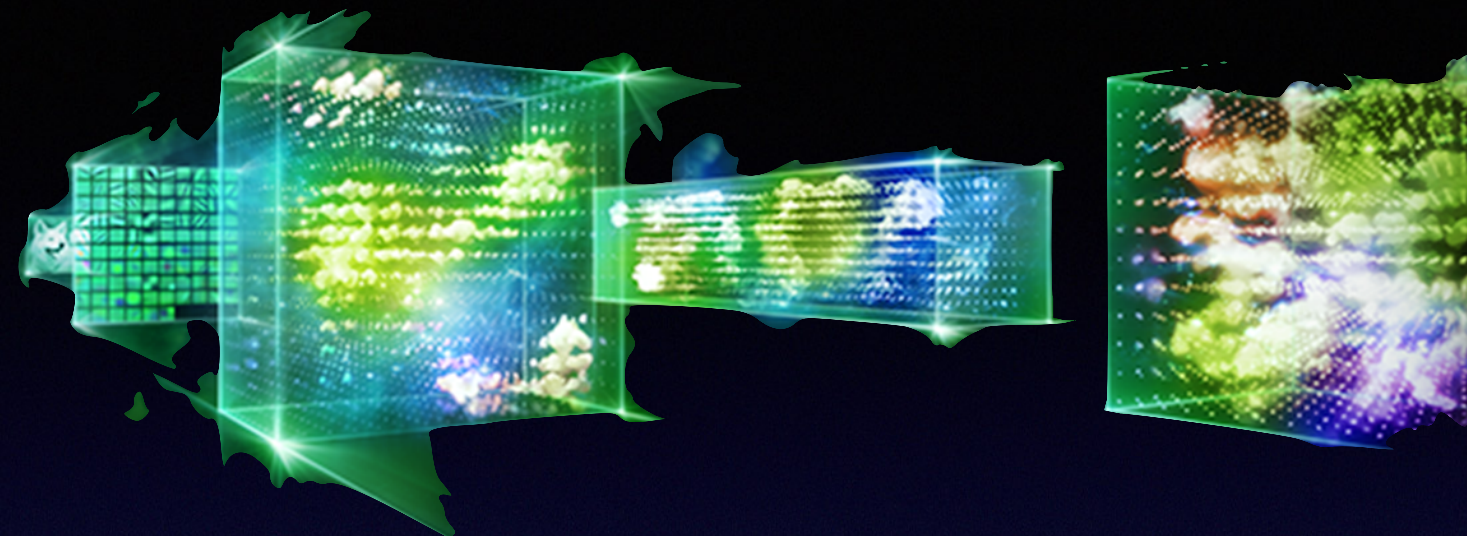
… enough challenges …



Solutions?

- **Path A**: “**traditional path**”
  - Hand-engineered reconstruction algorithms
- **Path B**: **machine learning**
  - Suited tool for a pattern recognition
  - “**Deep Learning**”
    - In particular… **Convolutional Neural Networks (CNNs)**
    - Scalable technique, generalizable to various tasks
    - Superb performance on image data analysis



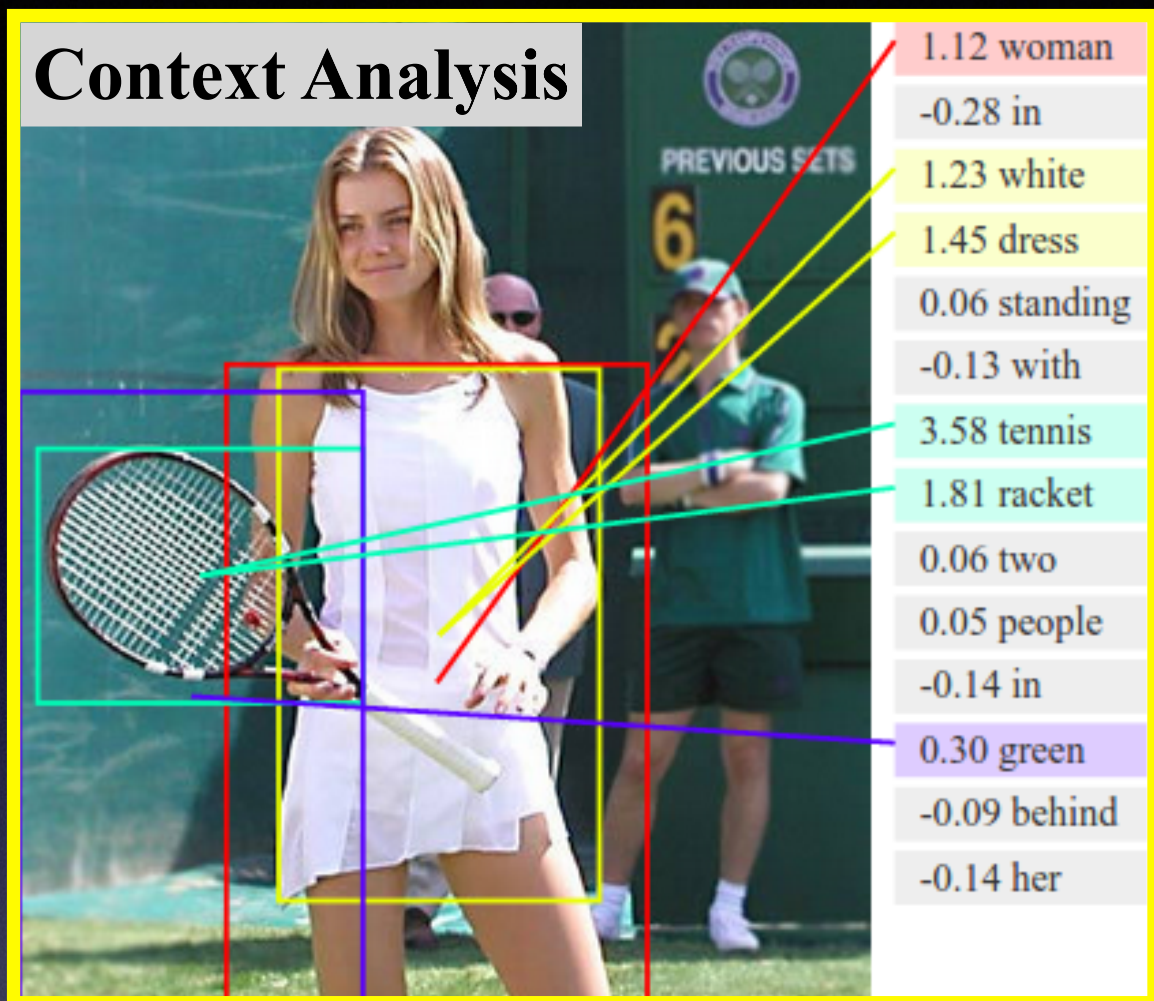


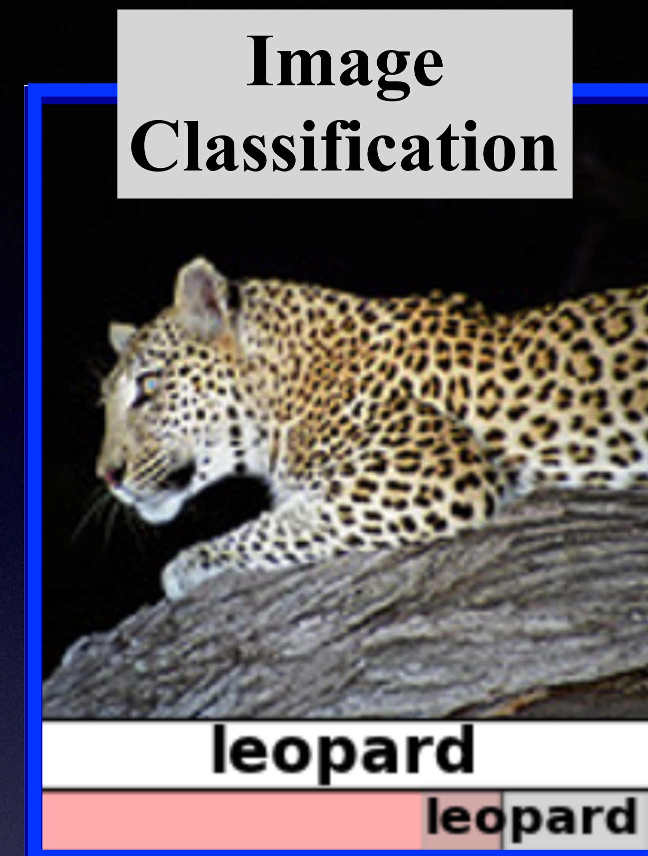
Convolutional Neural Networks for **LArTPC Analysis**

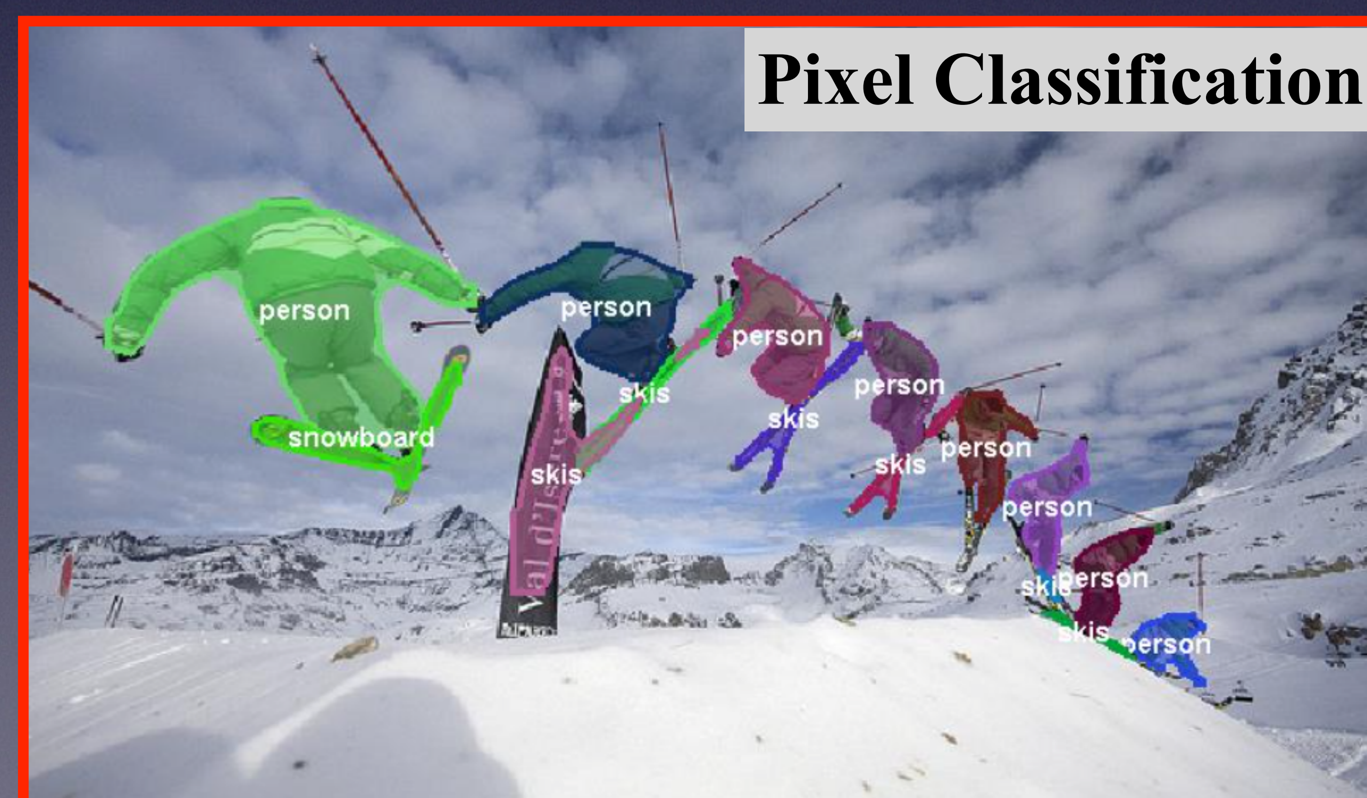
Outline

- Machine learning apps in MicroBooNE
- LArTPC image data + challenges
- Convolutional Neural Networks in MicroBooNE
- Summary

CNNs for Image Analysis





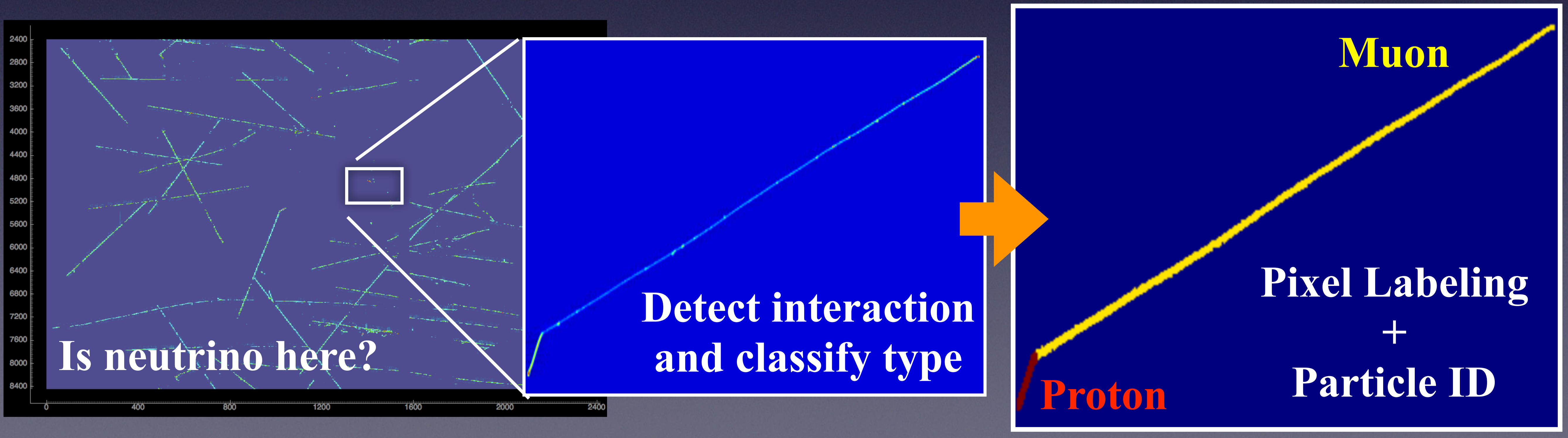


- Superb image analysis capabilities
- Trainable from raw data (large tensor)

CNN in MicroBooNE (I)

Applications in MicroBooNE

- **Event selection** (image classification)
- **Vertex finding** (object detection)
- **Clustering** (semantic segmentation)
- **Particle identification** (image classification)



$\nu_\mu + n \rightarrow \mu + p$

CNN in MicroBooNE (I)

Applications in MicroBooNE

- **Event selection** (image classification)
- **Vertex finding** (object detection)
- **Clustering** (semantic segmentation)
- **Particle identification** (image classification)

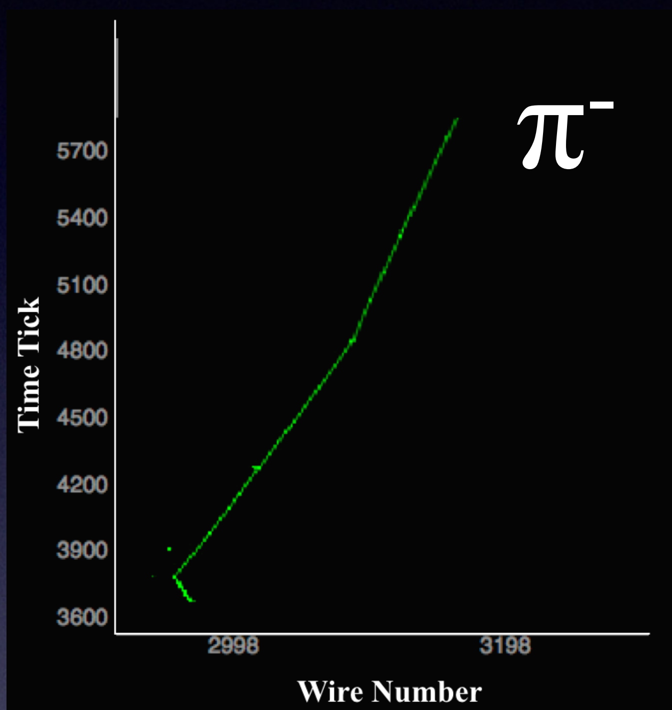
Highlights these CNNs in next slides

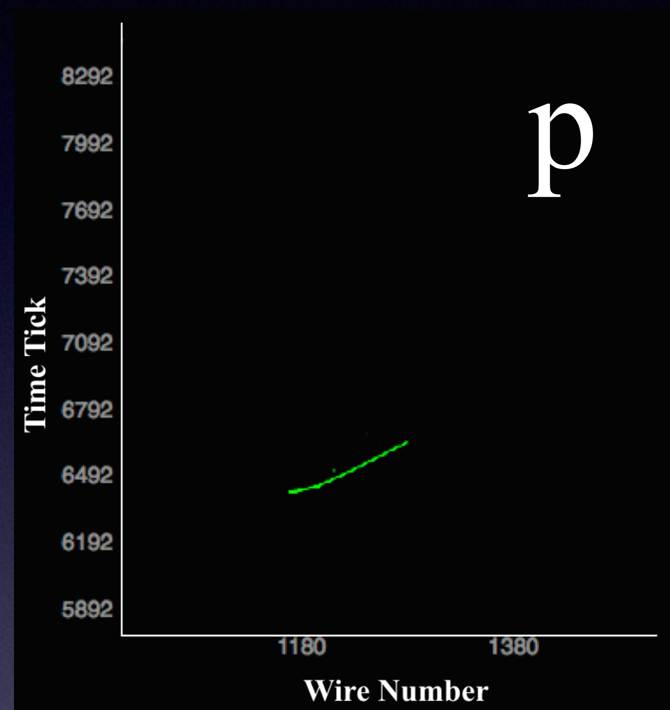
CNN in MicroBooNE (I)

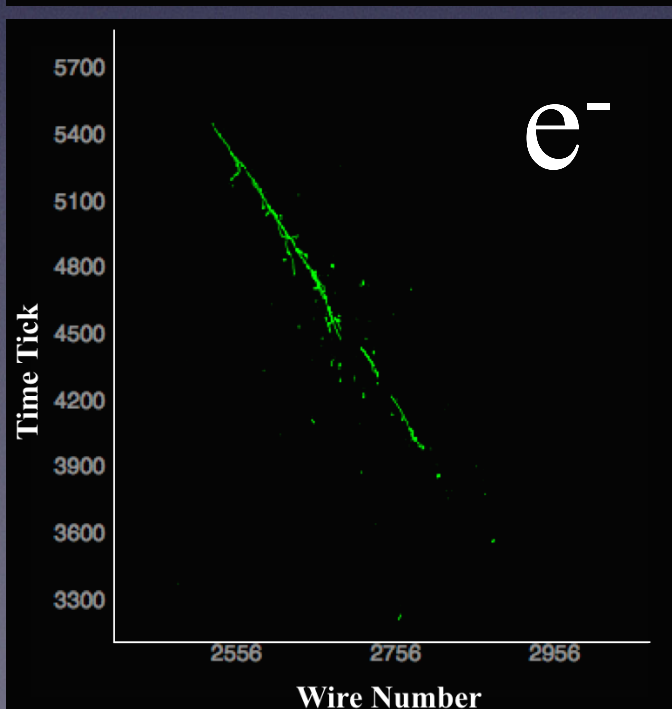
Particle identification

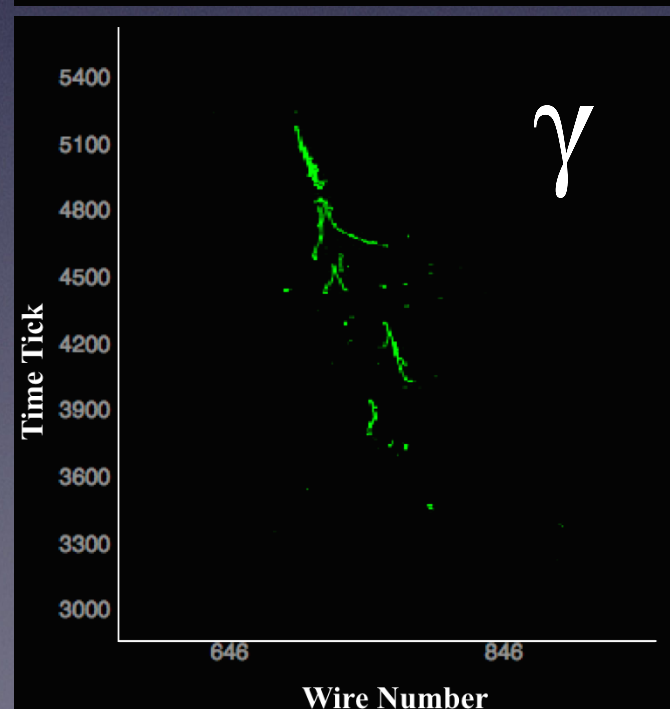
Trained a network to distinguish 5 particle types

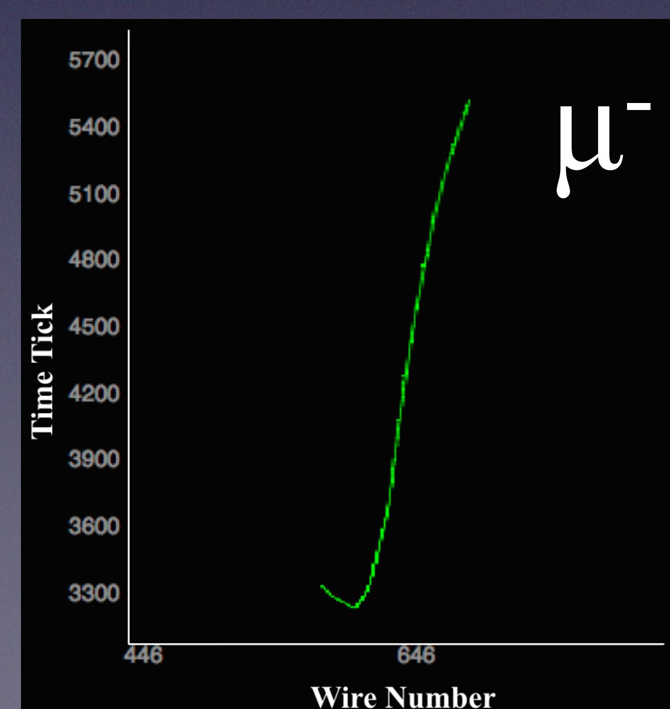
- **Simulated particles**
  - using 1 (collection) plane
- **Supervised training**
  - 22,000 images / type
- **Flat momentum dist.**
  - Uniform position
  - Isotropic [100, 1000] MeV/c











CNN in MicroBooNE (I)

Particle identification

Trained a network to distinguish 5 particle types

Particle	Efficiency	Mid-ID
e^-	0.778	γ ... 0.20
γ	0.834	e^- ... 0.15
μ^-	0.897	π^- ... 0.054
π^-	0.710	μ^- ... 0.226
proton	0.912	μ^- ... 0.046

JINST 10.1088/1748-9221

Further improvement?

- ~5 to 10% improvement by exploring network architectures
  - network width, effective depth
- Additional ~5% improvement by combining 3 planes using siamese architecture

Resource Usage

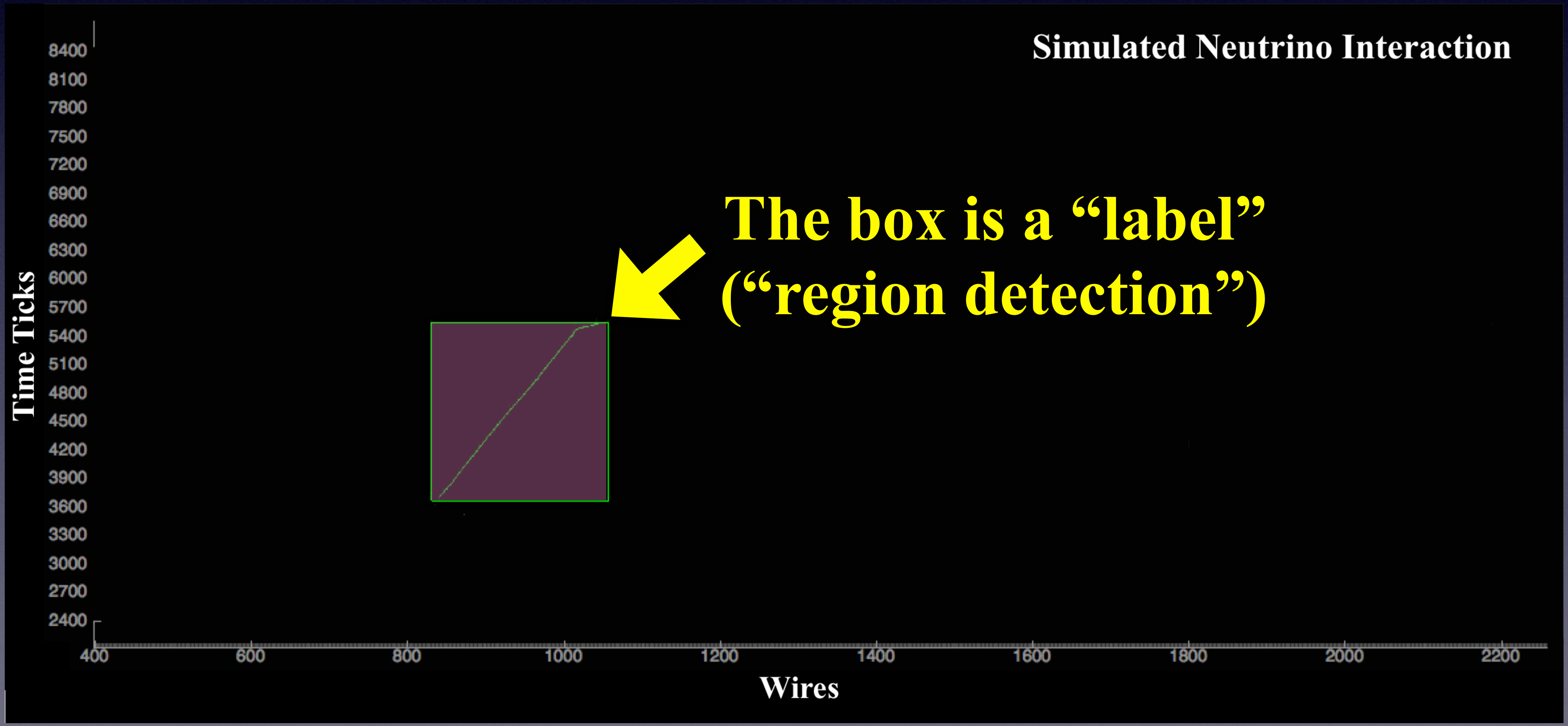
Architecture study include performance vs. speed!
Current architecture choice ~7 ms/image @ Titan X GPU)

CNN in MicroBooNE (II)

Event vertex detection

Trained a network to find neutrino interaction region

- Training sample uses simulated neutrino + cosmic data image
  - Supervised training using ≃100,000 collection plane images (1-plane)

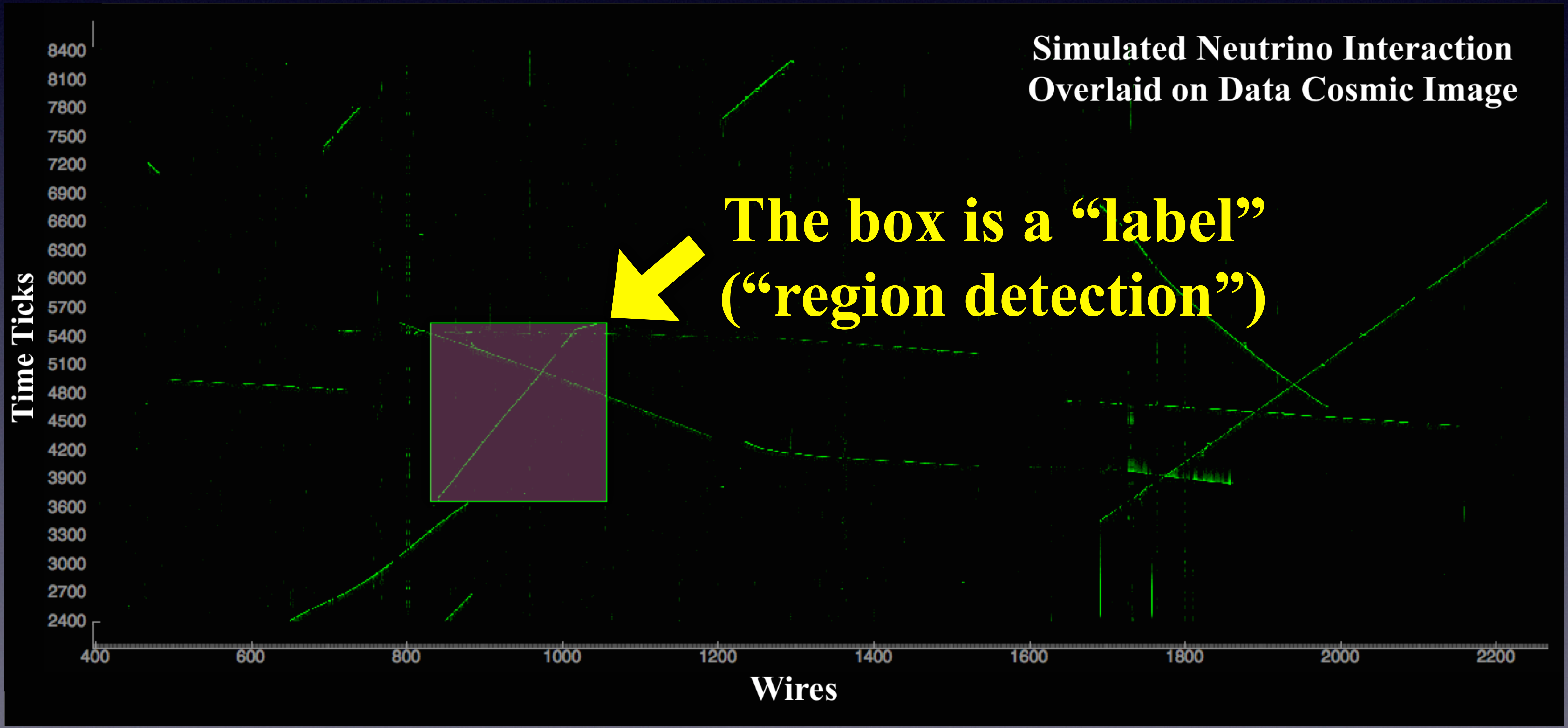


CNN in MicroBooNE (II)

Event vertex detection

Trained a network to find neutrino interaction region

- Training sample uses simulated neutrino + cosmic data image
  - Supervised training using ≃100,000 collection plane images (1-plane)

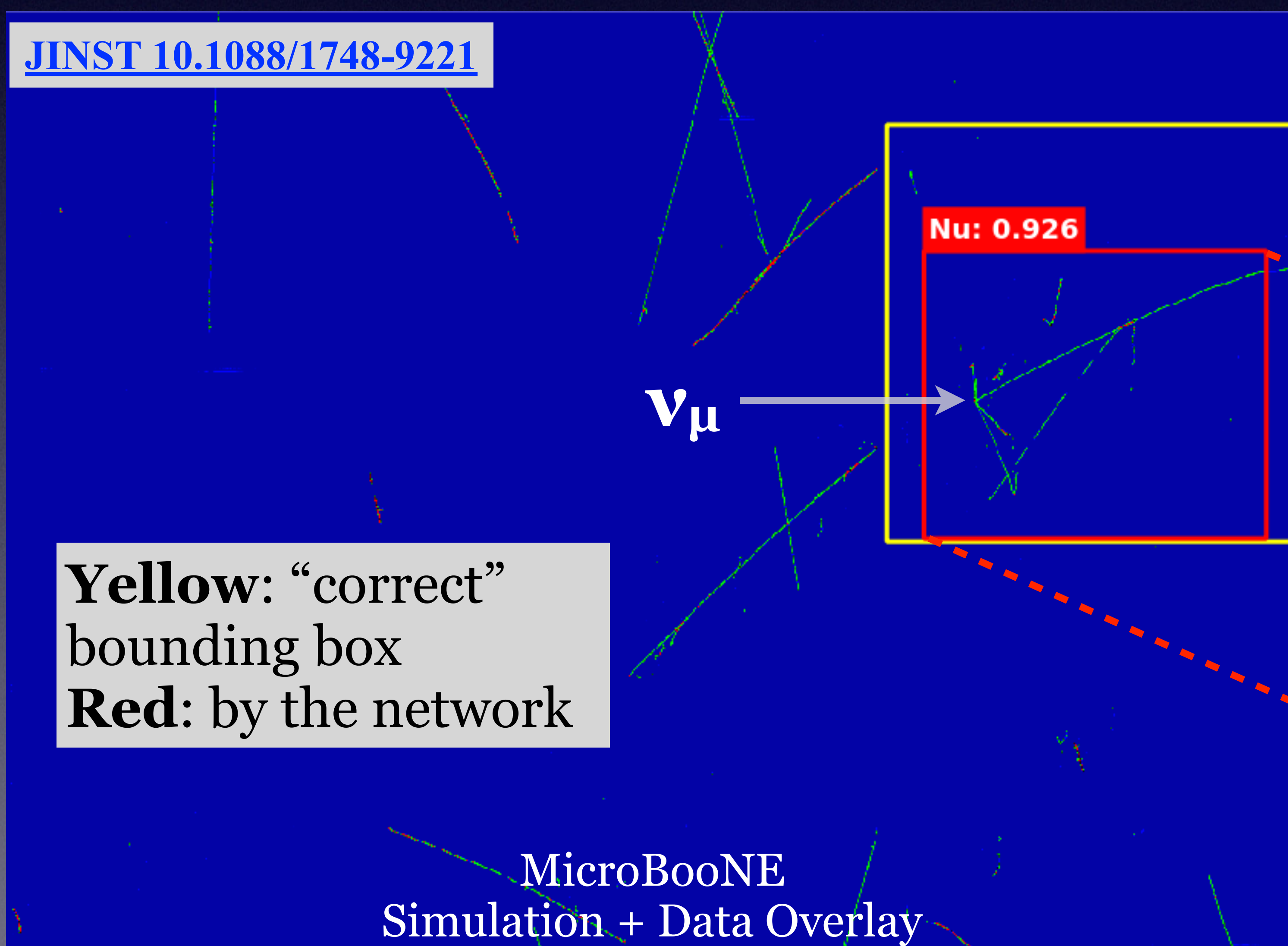


CNN in MicroBooNE (II)

Event vertex detection

Trained a network to find neutrino interaction region

- Training sample uses simulated neutrino + cosmic data image
  - Supervised training using ≃100,000 collection plane images (1-plane)



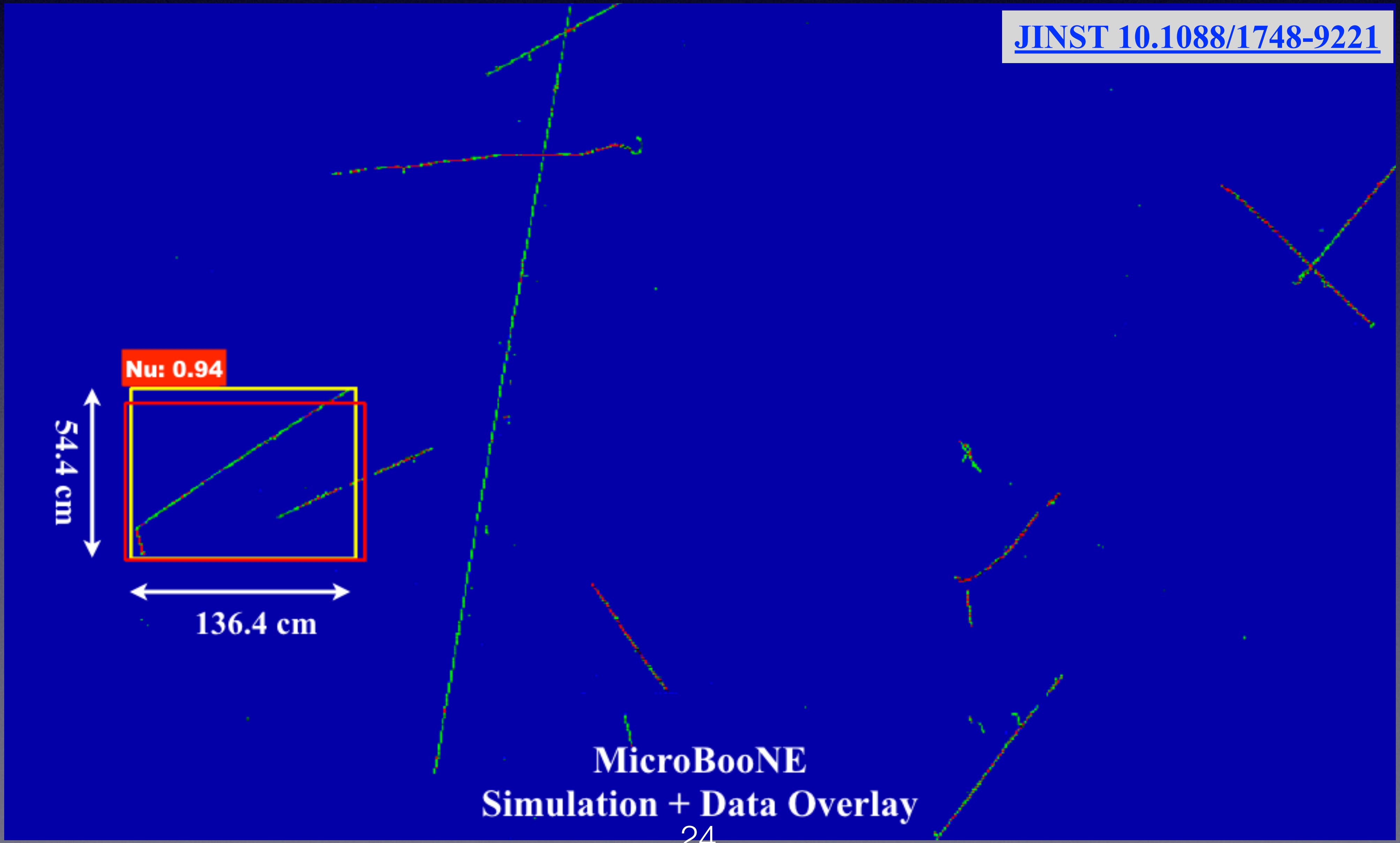
Network Output
≃ 2.6m (width) x 1 m (height)

CNN in MicroBooNE (II)

Event vertex detection

Trained a network to find neutrino interaction region

- Training sample uses simulated neutrino + cosmic data image
  - Supervised training using ≃100,000 collection plane images (1-plane)

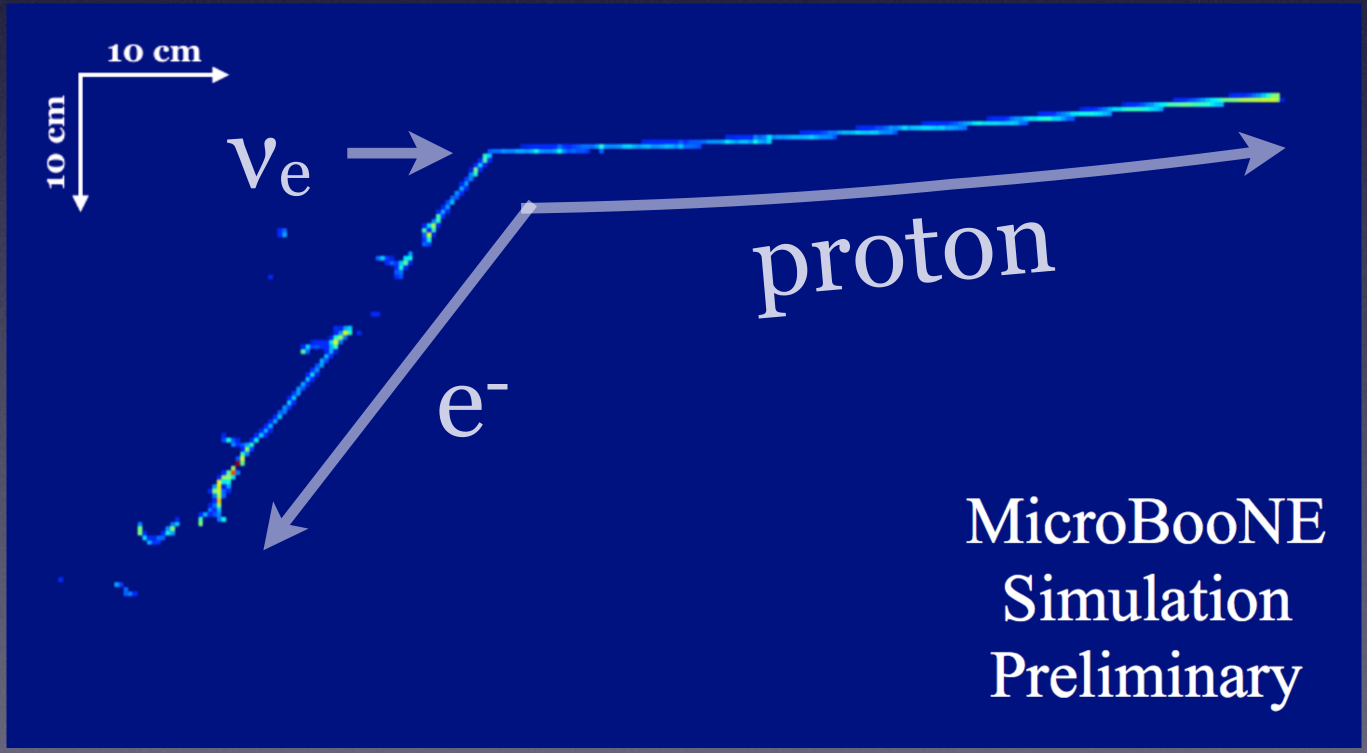


CNN in MicroBooNE (III)

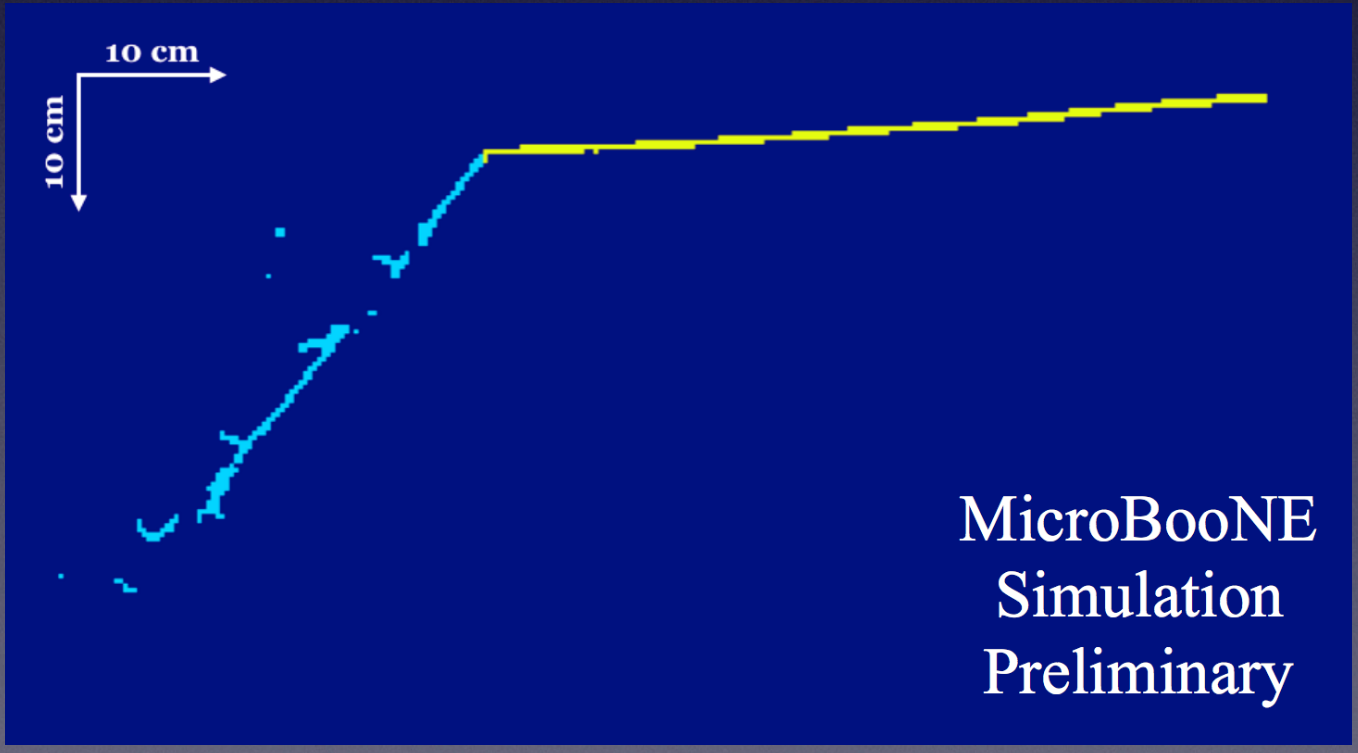
Particle clustering using a network

CNN designed to **segment pixels by predefined semantics**

- Can perform particle-wise pixel clustering
- On-going work
  - "track/shower" pixel labeling by the network (clustered by algorithm)
  - Custom training technique to improve performance on LArTPC image



ADC Image



Network Output

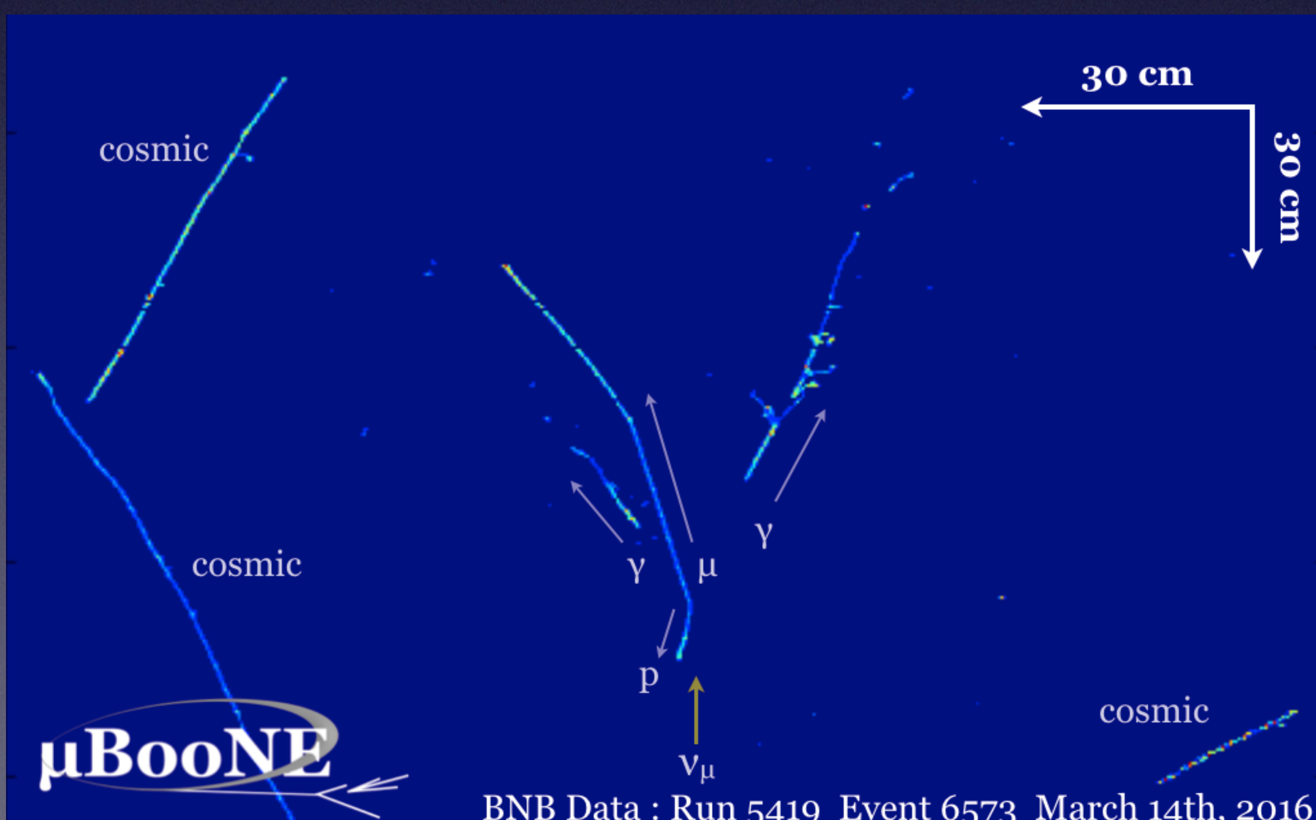
CNN in MicroBooNE (III)

Particle clustering using a network

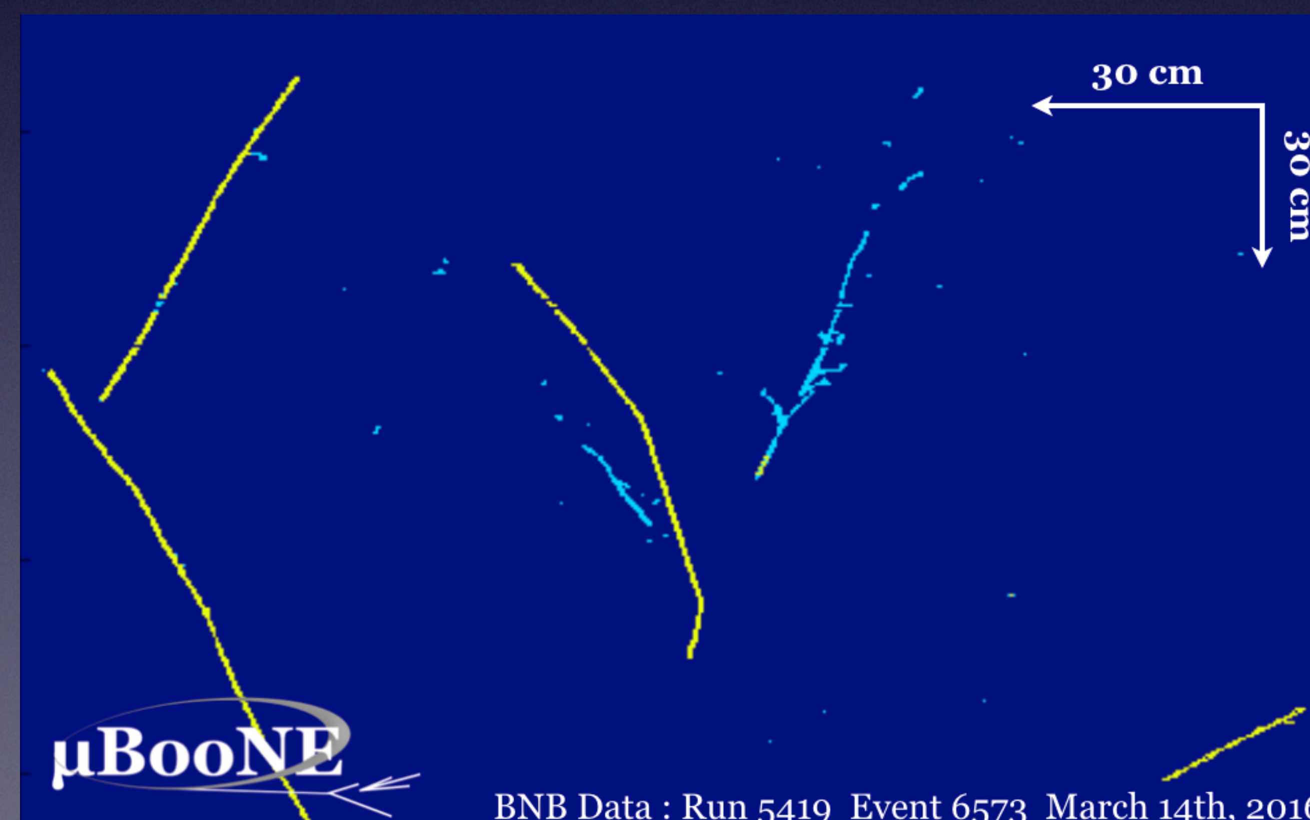
CNN designed to **segment pix** … **antics**

- Can perfor…
- On-going w…
  - "track/show… …abeling by the network (clustered by algorithm)
  - Custom training technique to improve performance on LArTPC image

Also making sure **the networks work on data** as we go!



ADC Image



Network Output

CNN in MicroBooNE (IV)

Optimize multiple tasks together (future project)

"Multi-task Network Cascade" can introduce task dependencies

- Allows to optimize the whole chain together

… sorry for my parenthood …

CNN in MicroBooNE (V)

Some studies published!

- **Event selection** (image classification)
- **Vertex finding** (object detection)
- **Clustering** (semantic segmentation)
- **Particle identification** (image classification)

Feel free to contact us for details!

MicroBooNE's 1st paper
JINST 10.1088/1748-9221
arXiv 1611.05531

Cornell University Library

arXiv.org > physics > arXiv:1611.05531

Physics > Instrumentation and Detectors

Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber

MicroBooNE collaboration: R. Acciarri, C. Adams, R. An, J. Asaadi, M. Auger, L. Bagby, B. Baller, G. Barr, M. Bass, F. Bay, M. Bishai, A. Blake, T. Bolton, L. Bugel, L. Camilleri, D. Caratelli, B. Carls, R. Castillo Fernandez, F. Cavanna, H. Chen, E. Church, D. Cianci, G. H. Collin, J. M. Conrad, M. Convery, J. I. Crespo-Anadón, M. Del Tutto, D. Devitt, S. Dytman, B. Eberly, A. Ereditato, L. Escudero Sanchez, J. Esquivel, B. T. Fleming, W. Foreman, A. P. Furmanski, G. T. Garvey, V. Genty, D. Goeldi, S. Gollapinni, N. Graf, E. Gramellini, H. Greenlee, R. Grosso, R. Guenette, A. Hackenburg, P. Hamilton, O. Hen, J. Hewes, C. Hill, J. Ho, G. Horton-Smith, C. James, J. Jan de Vries, C.-M. Jen, L. Jiang, R. A. Johnson, B. J. P. Jones, J. Joshi, H. Jostlein, D. Kaleko, G. Karagiorgi, W. Ketchum, et al. (75 additional authors not shown)

(Submitted on 17 Nov 2016)

We present several studies of convolutional neural networks applied to data coming from the MicroBooNE detector, a liquid argon time projection chamber (LArTPC). The algorithms studied include the classification of single particle images, the localization of single particle and neutrino interactions in an image, and the detection of a simulated neutrino event overlaid with cosmic ray backgrounds taken from real detector data. These studies demonstrate the potential of convolutional neural networks for particle identification or event detection on simulated neutrino interactions. We also address technical issues that arise when applying this technique to data from a large LArTPC at or near ground level.

CNN in MicroBooNE (VI)

LArTPC "real data" sample

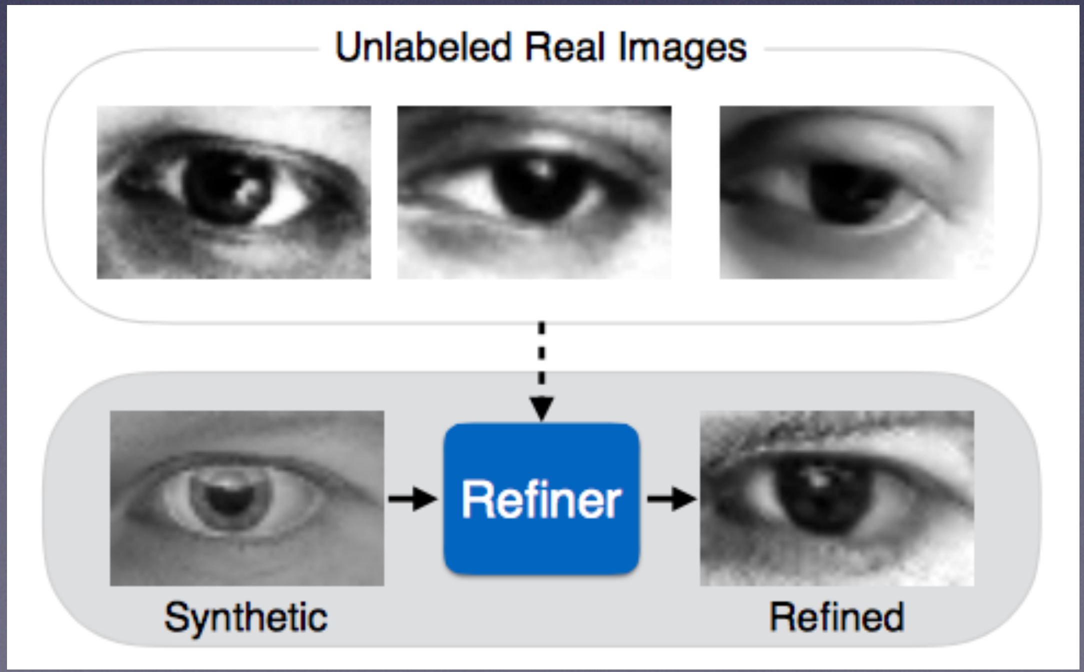
MicroBooNE provides LArTPC image data from the real detector

"Labeled" image database

- "neutrino event or not", "region of interaction", "particle type", etc…
- **Key to accelerate the technique R&D** (similar to ILSVRC in CV)

Generative Adversarial Networks (GANs)

- New technique, great intellectual interest in the field (both CV and HEP)
- Possibility to **learn generic image features from real data**
  - Semi-supervised or unsupervised learning
- Example applications
  - LArTPC image generator
  - **MC image "encoder"**
    - Refine "MC images" into "data images"



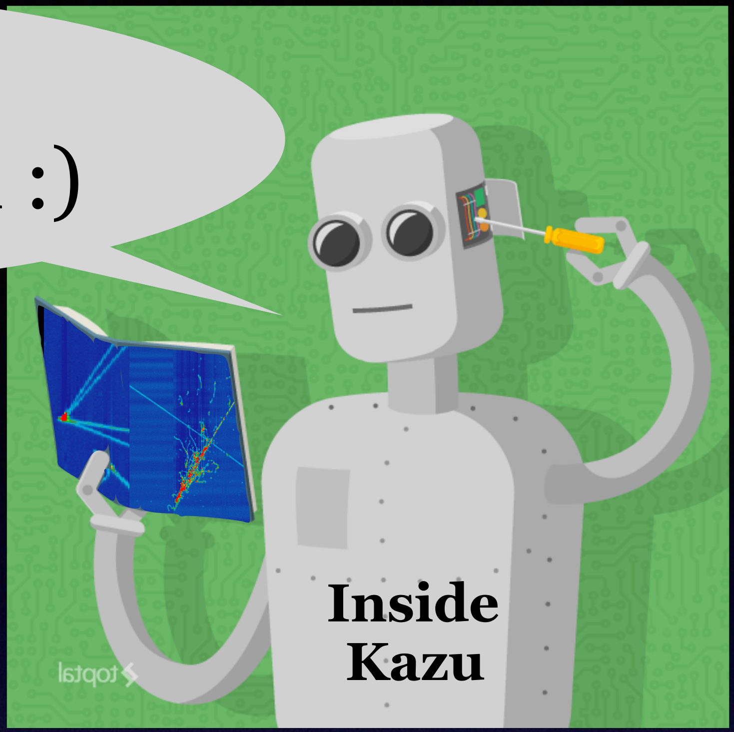
"**encoder**" for human eye illustration
by Apple research team
arXiv:1612.07828

… wrapping up …

Outline

- Machine learning apps in MicroBooNE
- LArTPC image data + challenges
- Convolutional Neural Networks in MicroBooNE
- **Summary**

Thank you!
for your attention :)



Take Away Messages

1. LArTPCs are **high precision detectors**

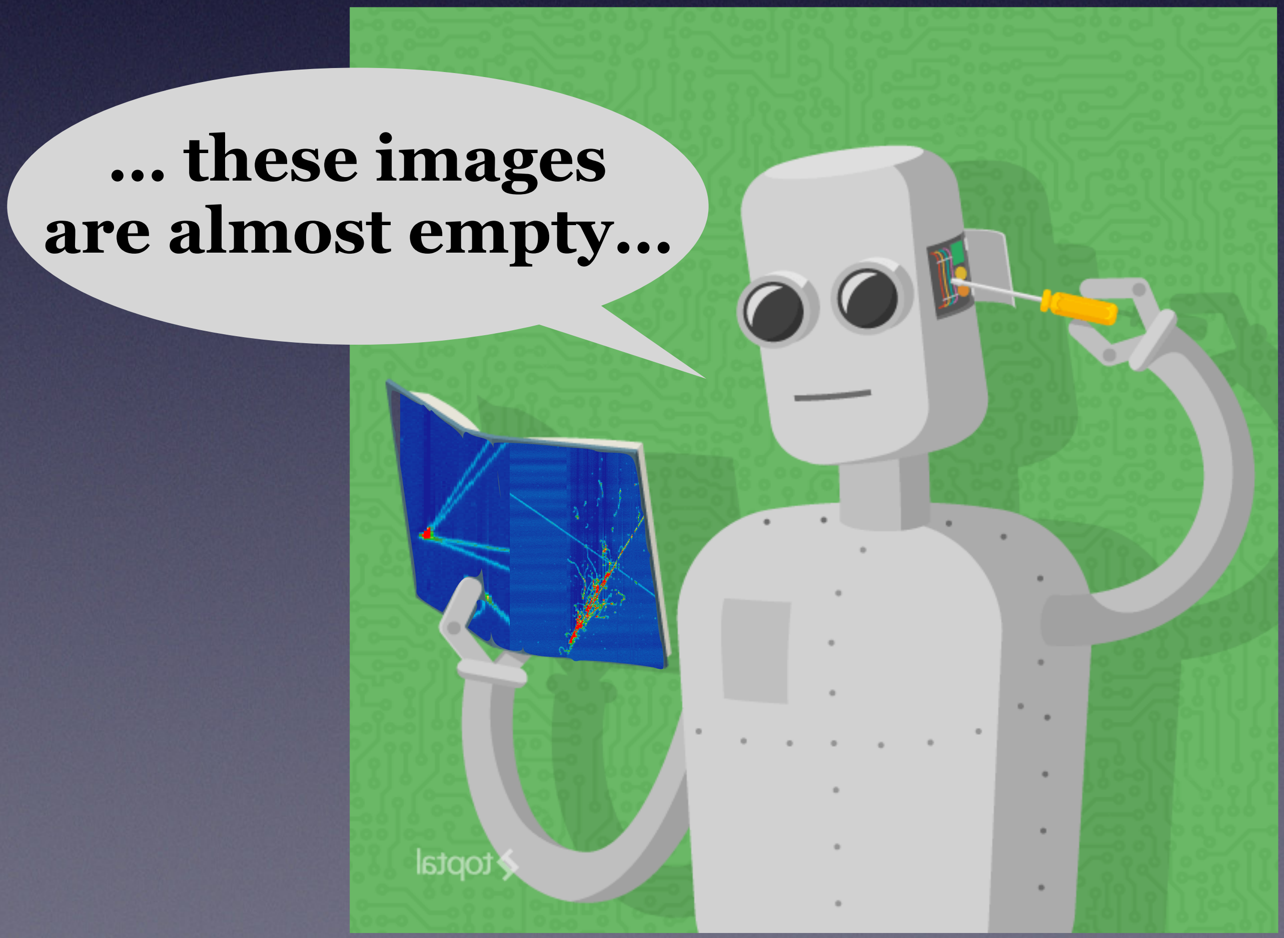
2. LArTPCs need **advanced pattern recognition algorithms**

3. MicroBooNE utilize **CNN for reconstruction/analysis**

4. MicroBooNE provides **unique opportunity to study real large scale LArTPC image data,** and **important challenges to overcome for future LArTPC detectors**

Back up

CNN for LArTPC
Image Analysis

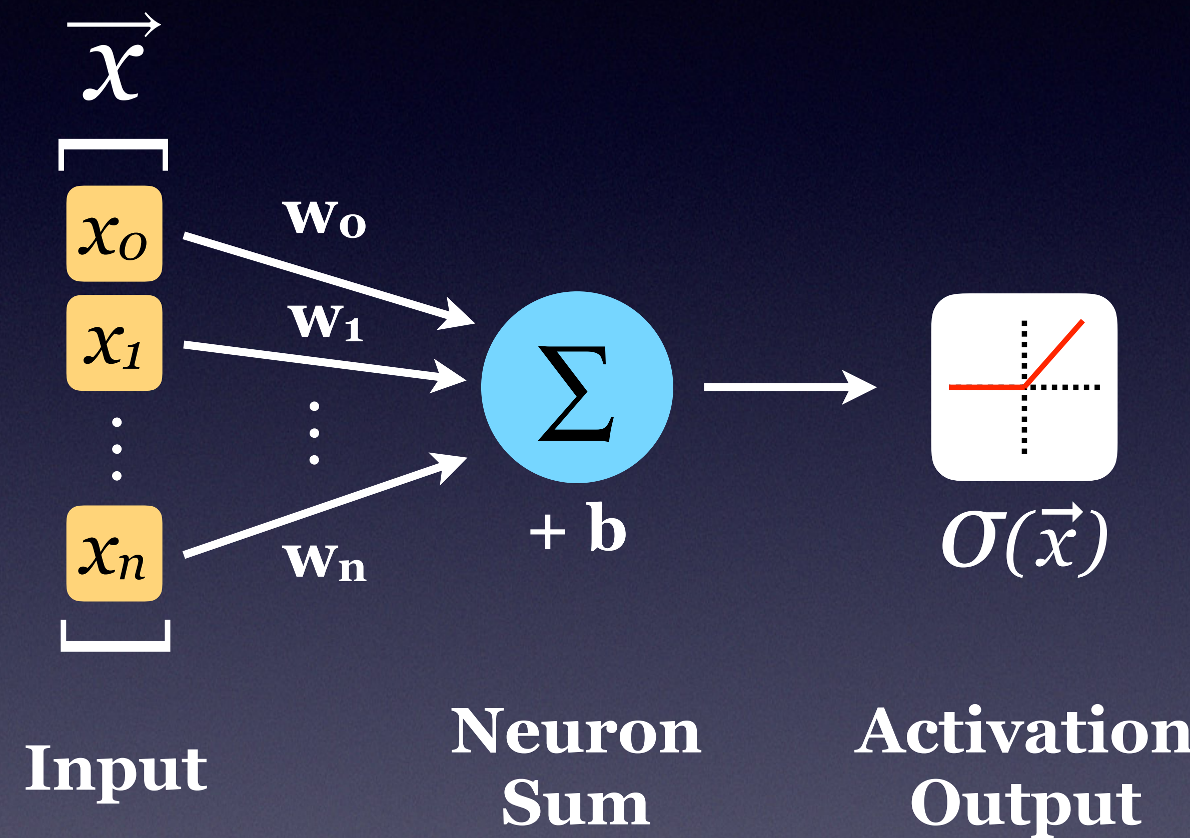


Introduction to CNNs (II)

Background: Neural Net

The basic unit of a neural net is the ***perceptron*** (loosely based on a real neuron)

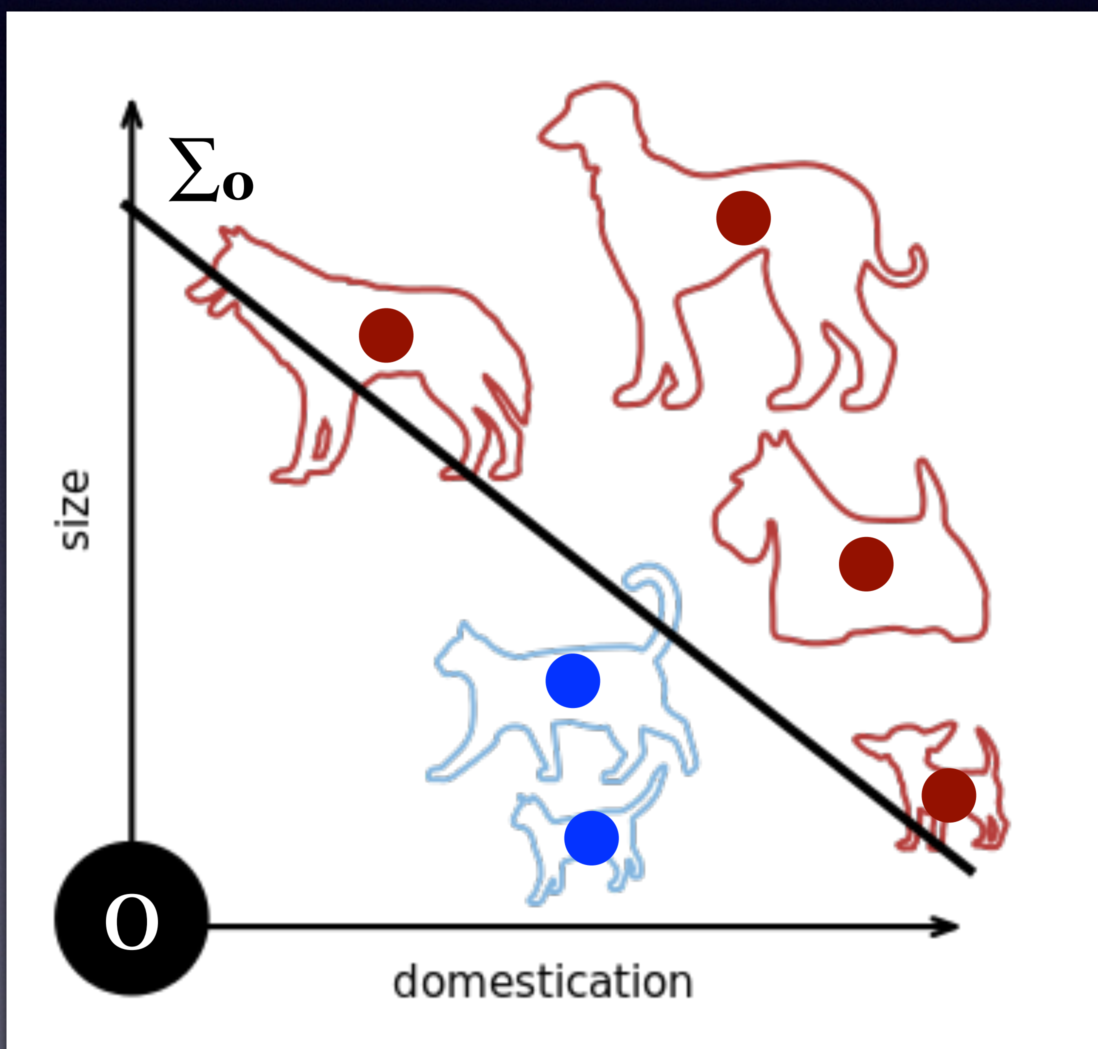
Takes in a vector of inputs (*x*). Commonly inputs are summed with weights (*w*) and offset (*b*) then run through activation.



$$\sigma(\vec{x}) = \begin{cases} \vec{w}_i \cdot \vec{x} + b_i & \vec{w}_i \cdot \vec{x} + b_i \geq 0 \\ 0 & \vec{w}_i \cdot \vec{x} + b_i < 0. \end{cases}$$

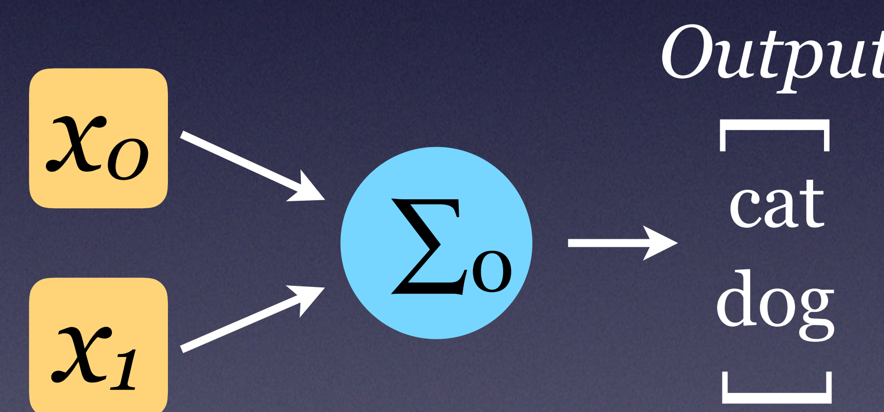
Introduction to CNNs (II)
Perceptron 2D Classification

Imagine using two features to separate cats and dogs



from [wikipedia]

$$\sigma(\vec{x}) = \begin{cases} \vec{w}_i \cdot \vec{x} + b_i & \vec{w}_i \cdot \vec{x} + b_i \geq 0 \\ 0 & \vec{w}_i \cdot \vec{x} + b_i < 0. \end{cases}$$

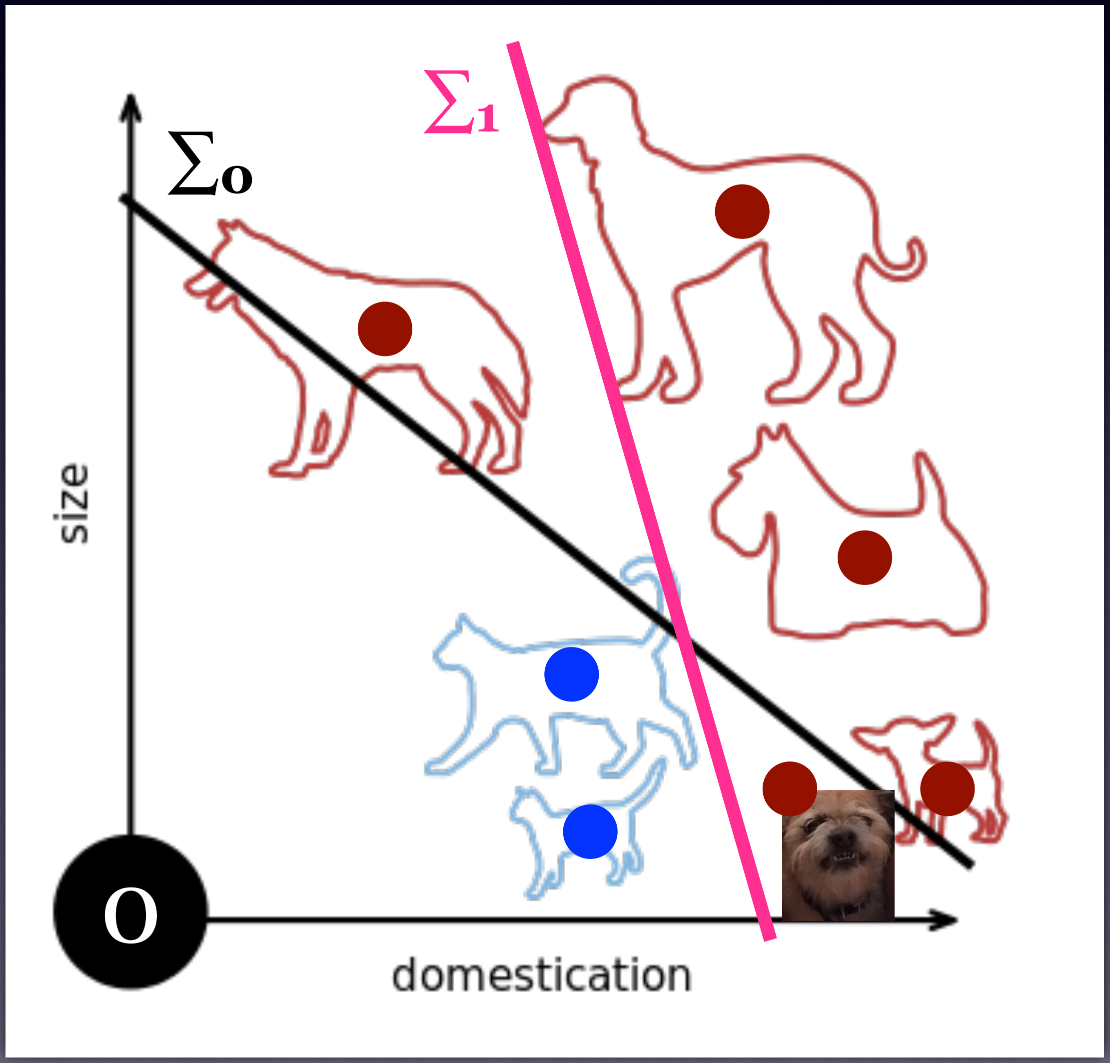


By picking a value for w and b, we define a boundary between the two sets of data

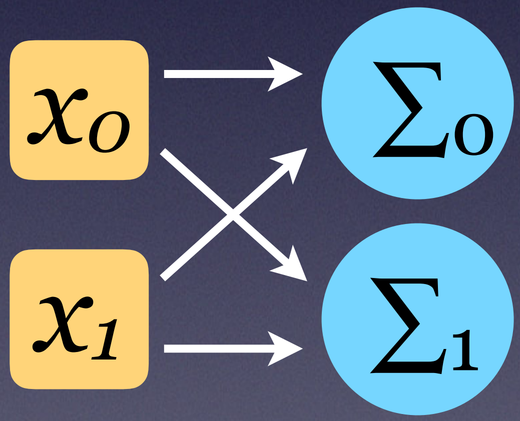
Introduction to CNNs (II)

Perceptron 2D Classification

Maybe we need to do better: assume new data point
(My friend's dog — small but not as well behaved)



from wikipedia

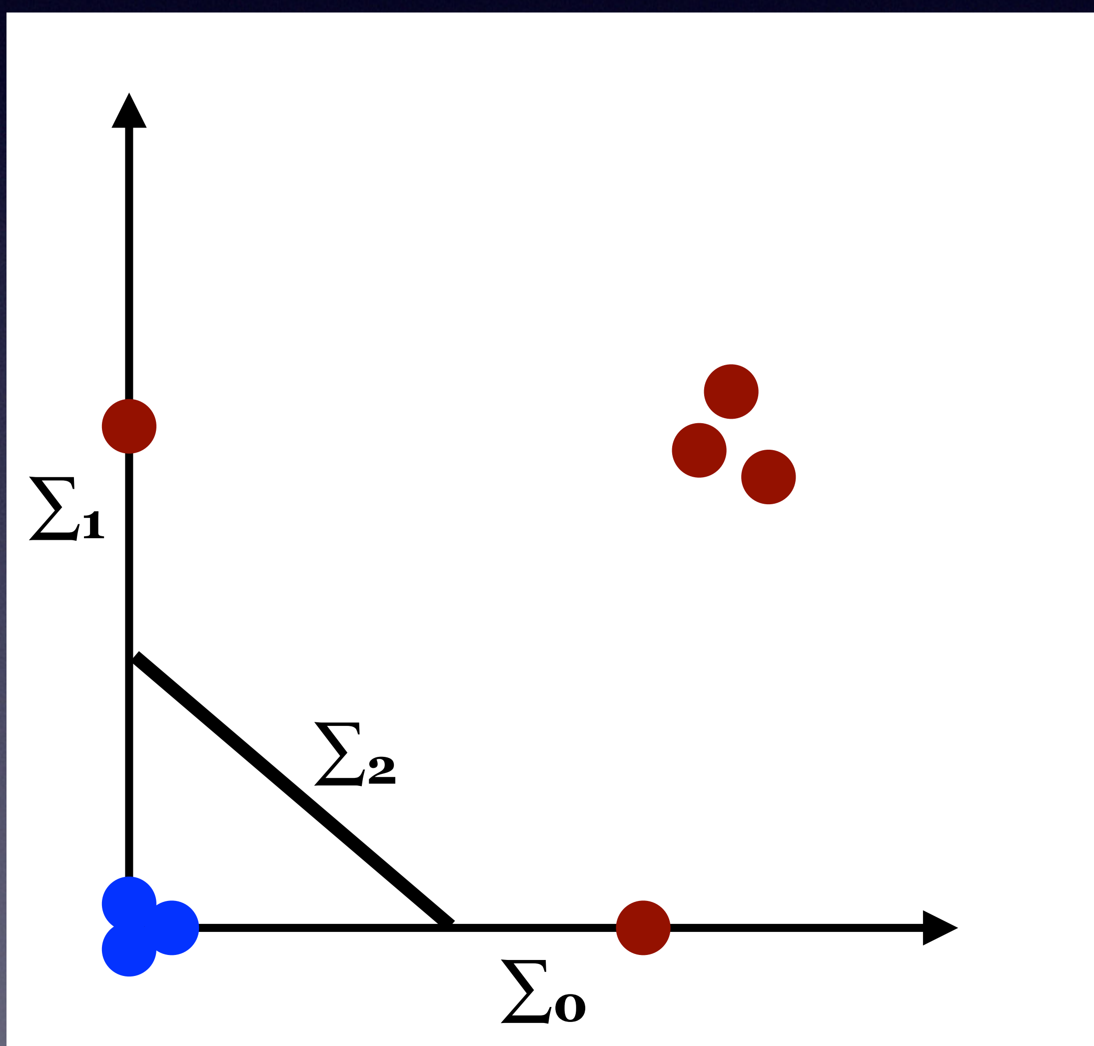


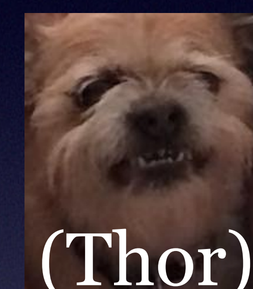
We can add another perceptron to help classify better

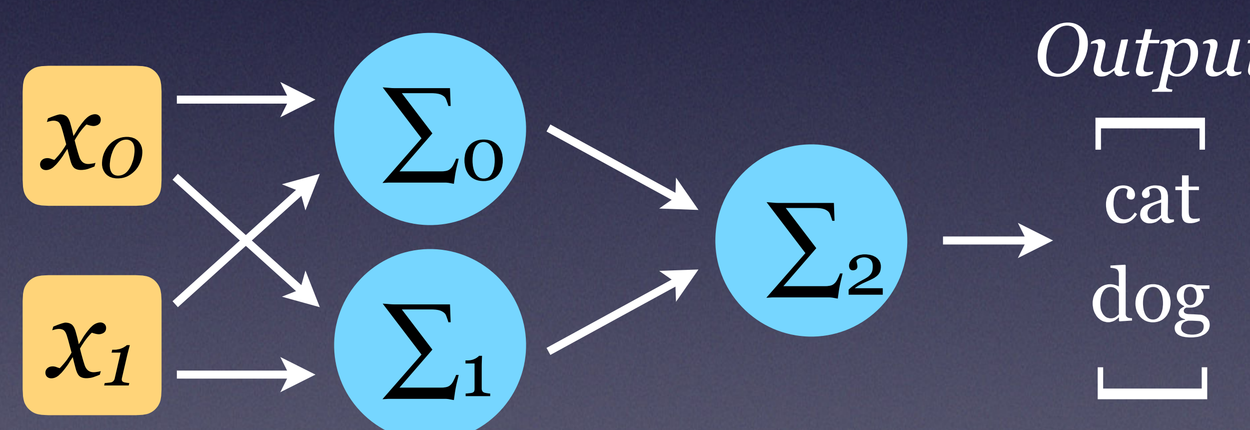
Introduction to CNNs (II)

Perceptron 2D Classification

Maybe we need to do better: assume new data point
(My friend's dog — small but not as well behaved)





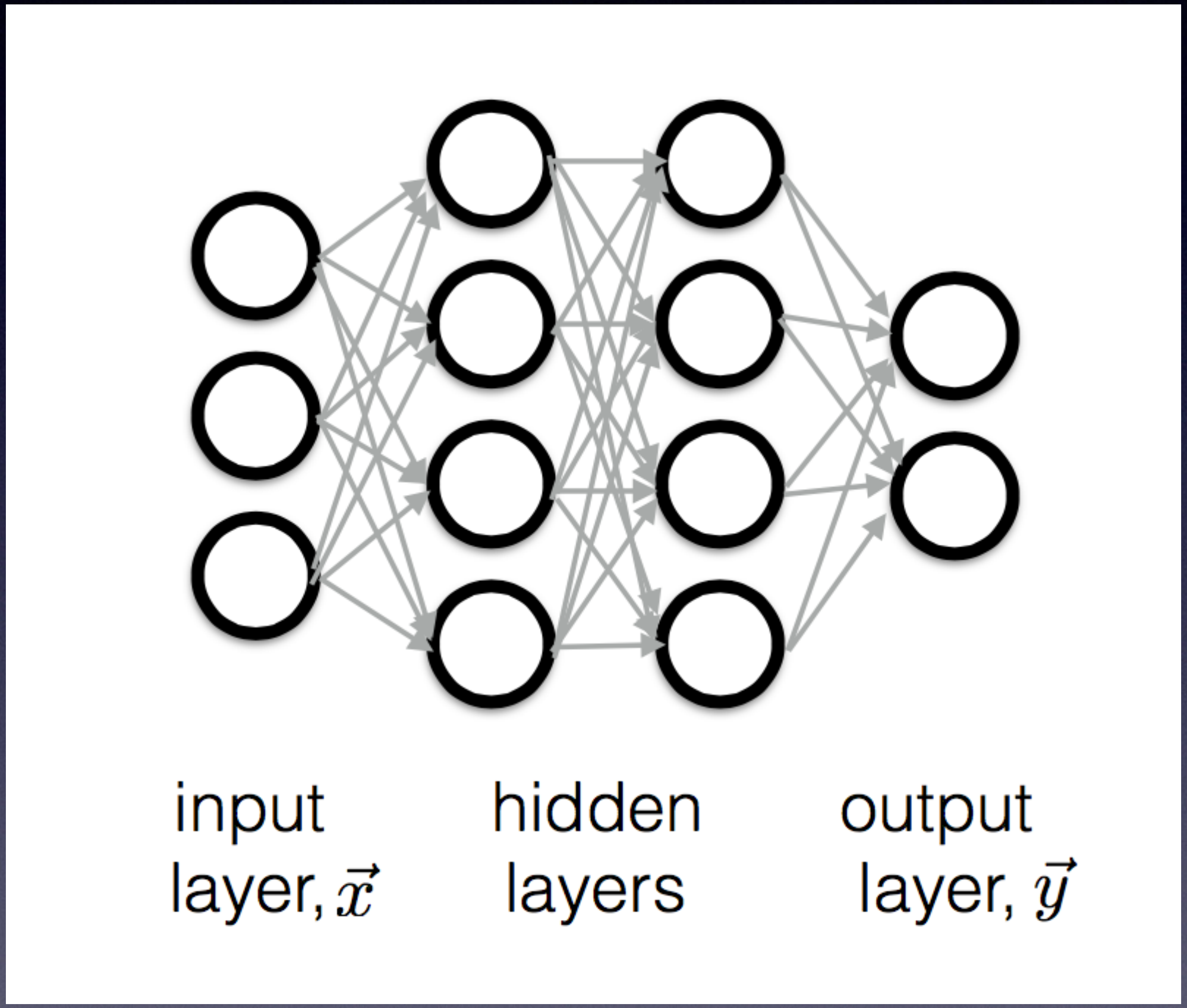


Another layer can classify based on preceding feature layer output

Introduction to CNNs (III)

"Traditional neural net" in HEP

Fully-Connected Multi-Layer Perceptrons



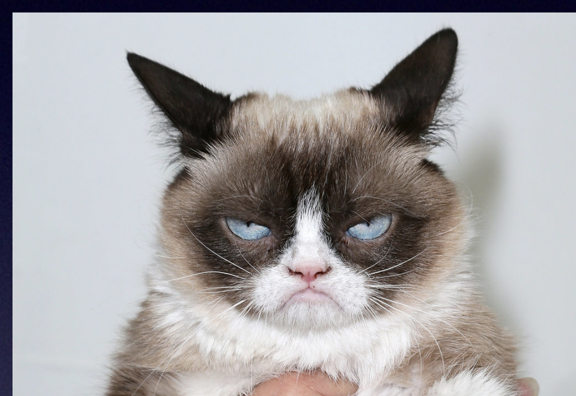
A traditional neural network consists of a stack of layers of such neurons where each neuron is ***fully connected*** to other neurons of the neighbor layers

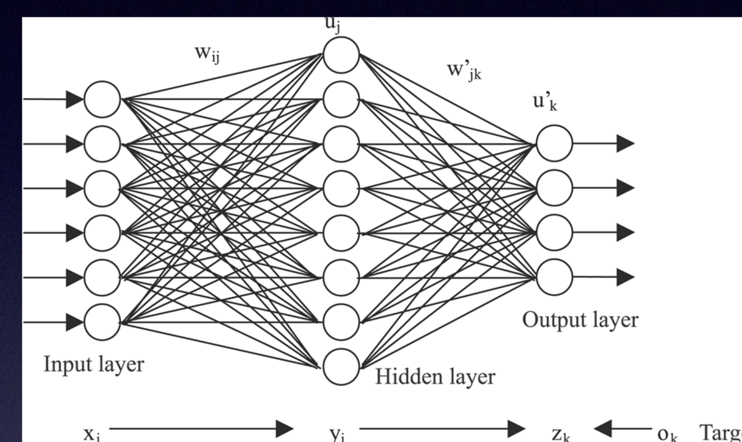
Introduction to CNNs (III)

"Traditional neural net" in HEP

Problems with it...

Feed in entire image

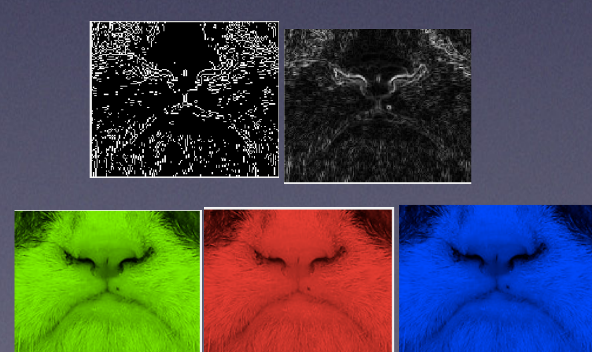


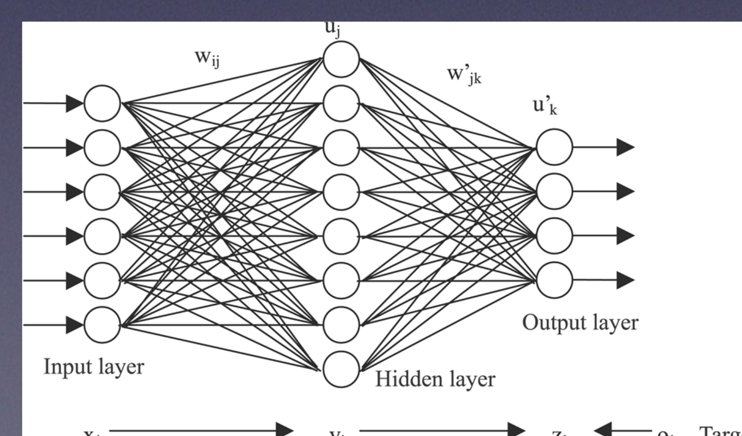


Cat?

Problem: scalability

Use pre-determined features



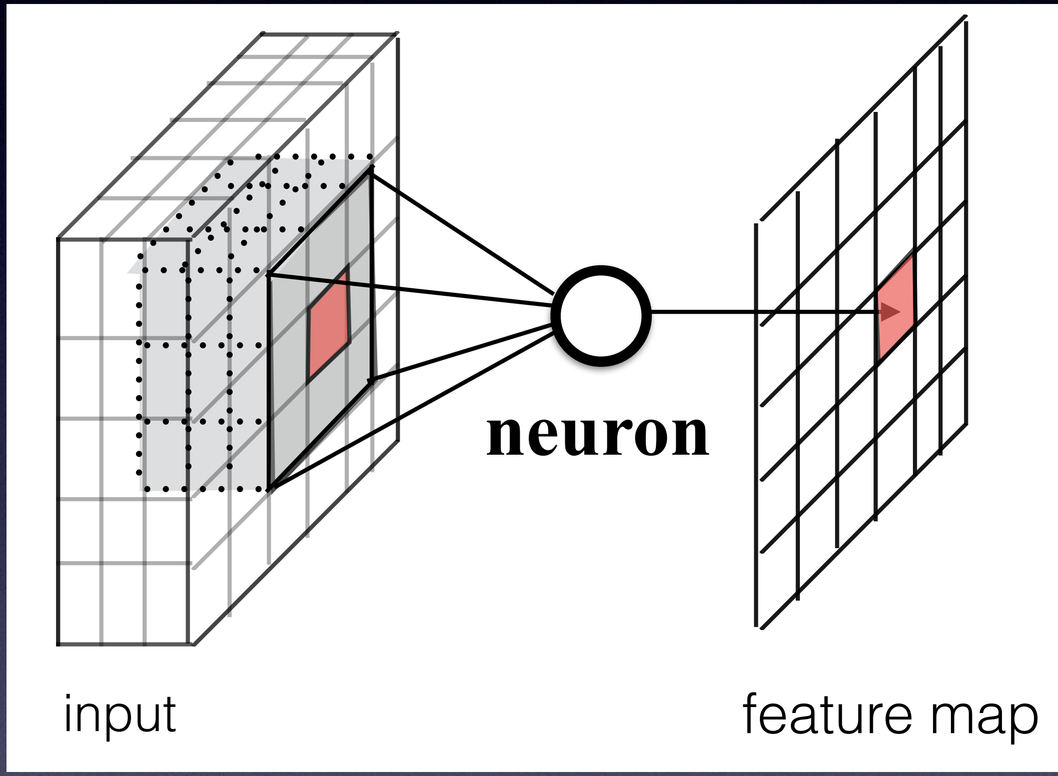


Cat?

Problem: generalization

Introduction to CNNs (III)

CNN introduce a ***limitation*** by forcing the network to look at only local, translation invariant features



$$f_{i,j}(X) = \sigma\left(W_i \cdot X_j + b_i\right),$$

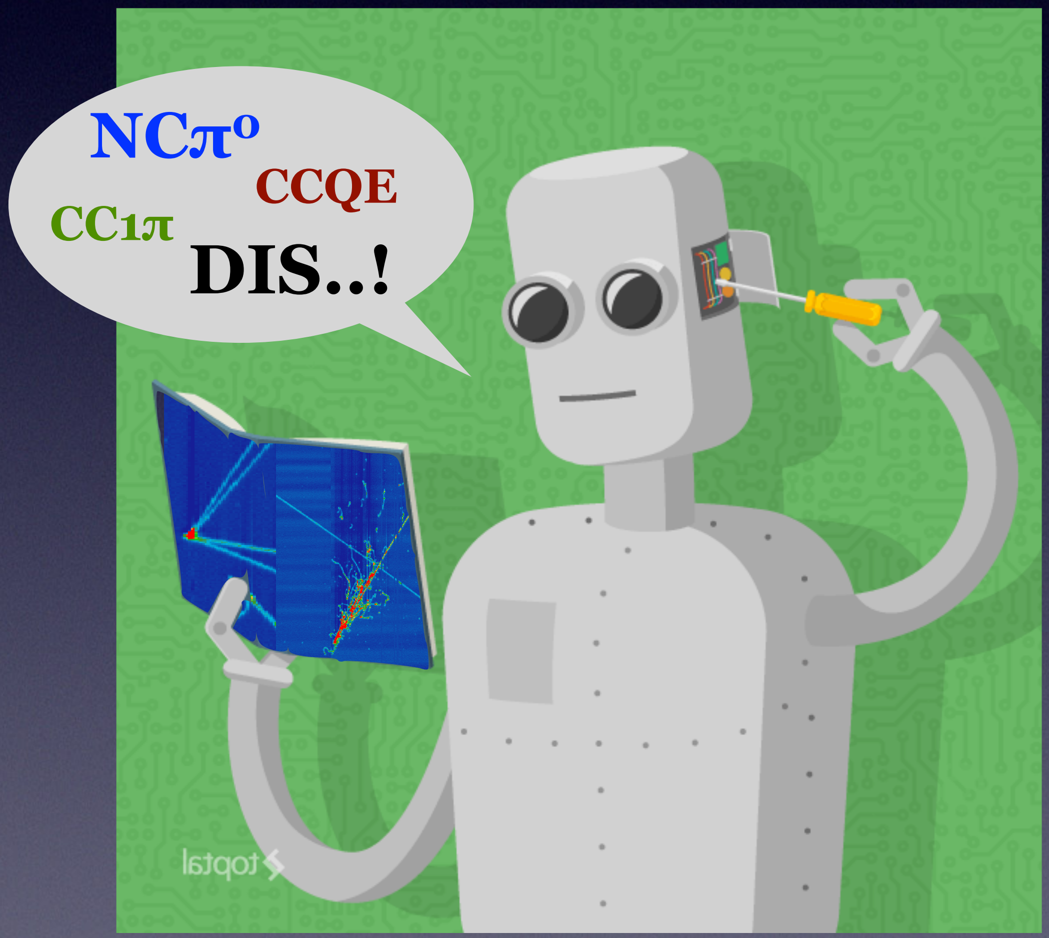
Activation of a neuron depends on the element-wise product of 3D weight tensor with 3D input data and a bias term

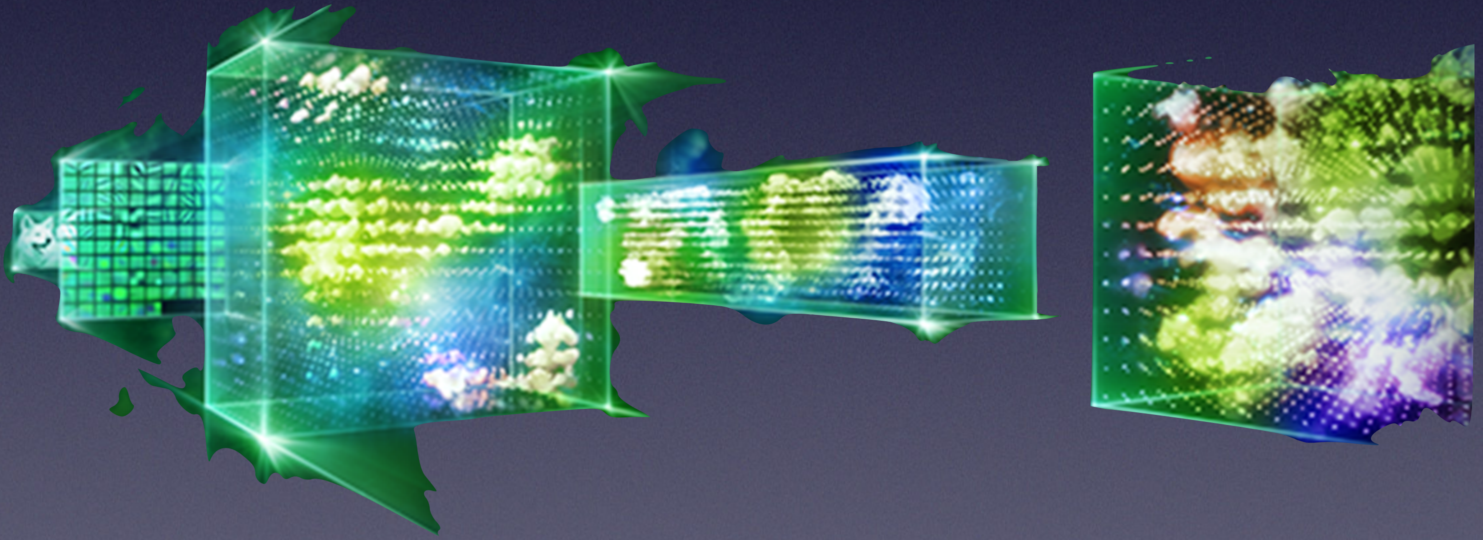
- Translate over 2D space to process the whole input
- Neuron **learns translation-invariant features**
- Applicable for a “homogeneous” detector like LArTPC

Want more details?
Feel free to ask me later!

Track/Shower Pixel Labeling

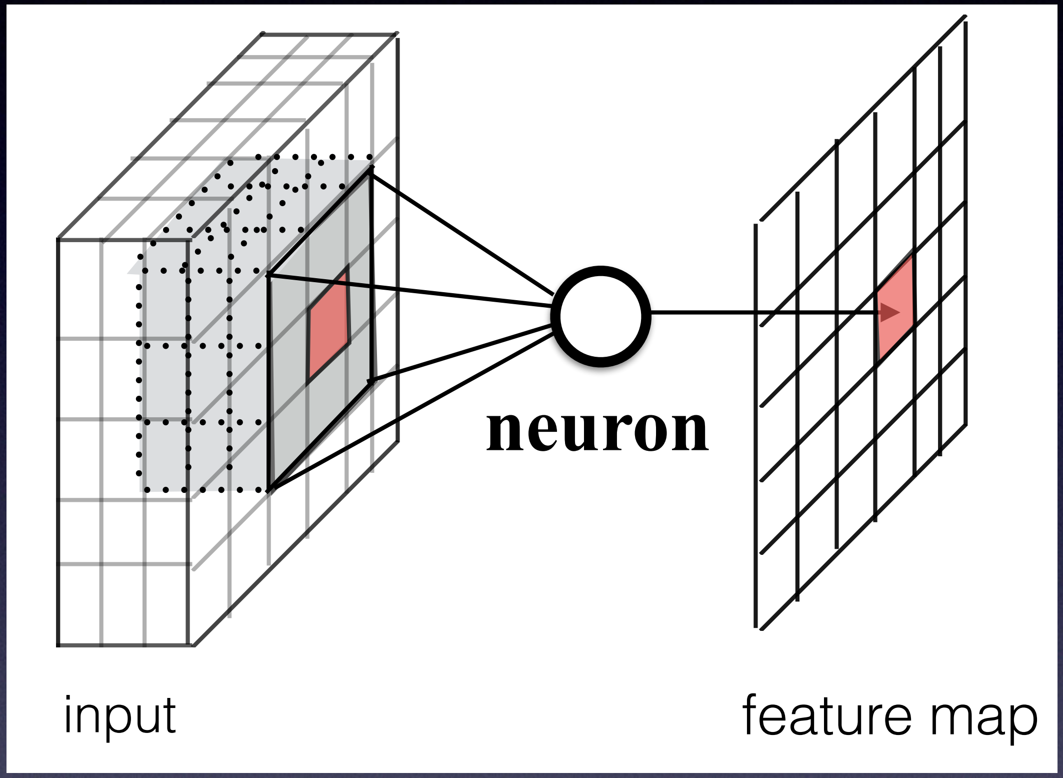
~ How Does SSNet Work? ~





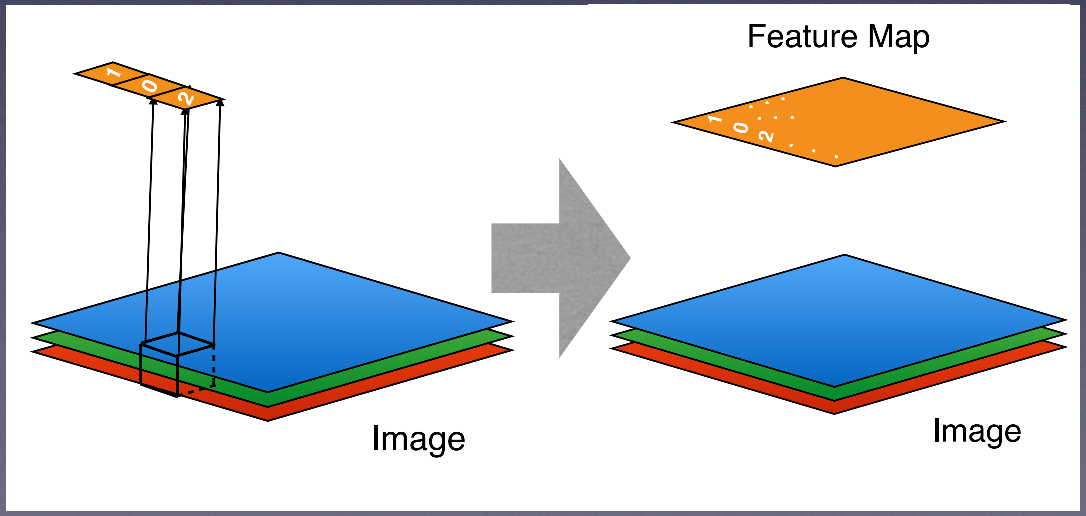
Quick Recap on CNN

CNN is a neural network formed with multiple convolution layers of neurons



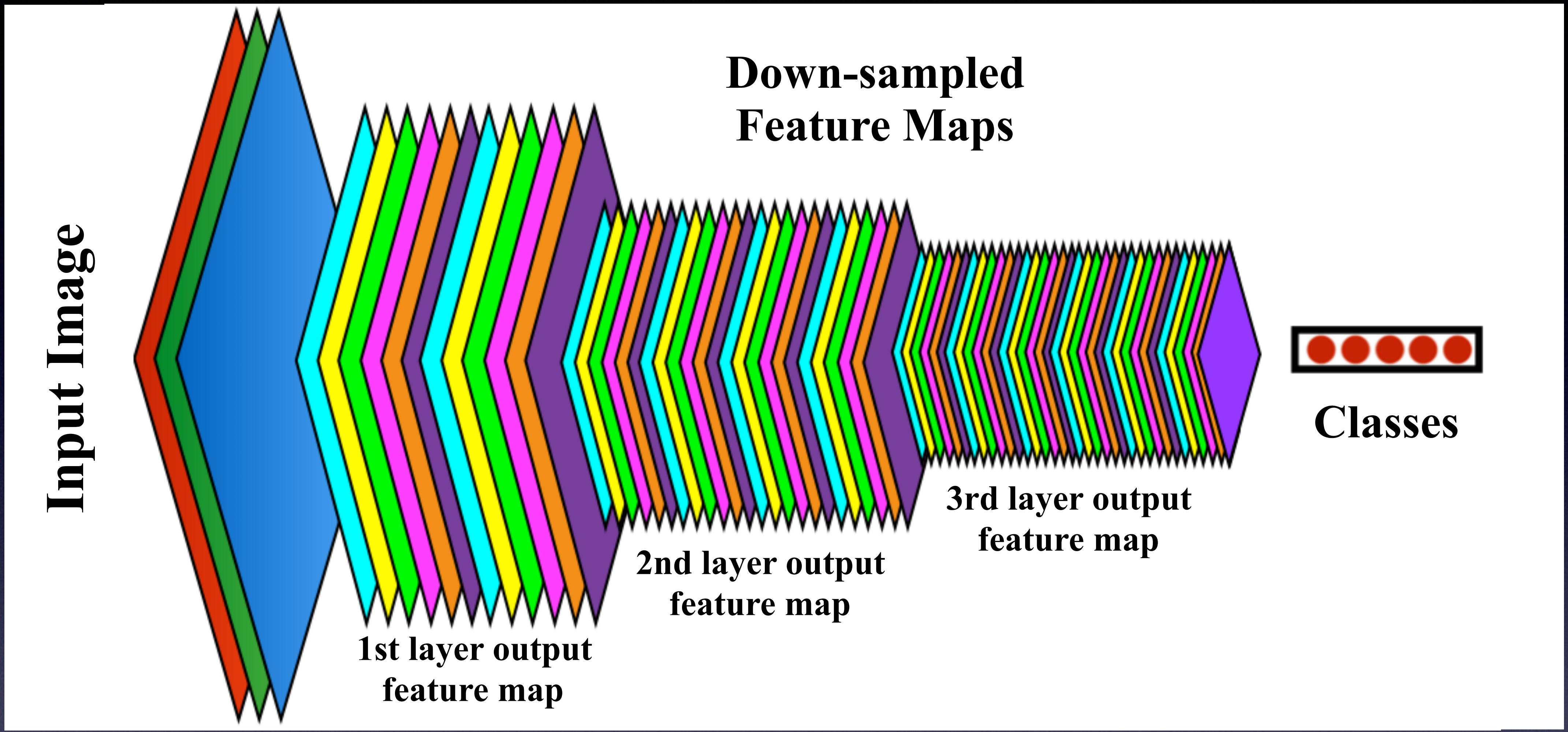
$$f_{i,j}(X) = \sigma\left(W_i \cdot X_j + b_i\right),$$

Activation of a neuron depends on the element-wise product of 3D weight tensor with 3D input data and a bias term



Each **filter** (neuron) translates over 2D space to process the whole input, producing a "***feature map***".

Quick Recap on CNN



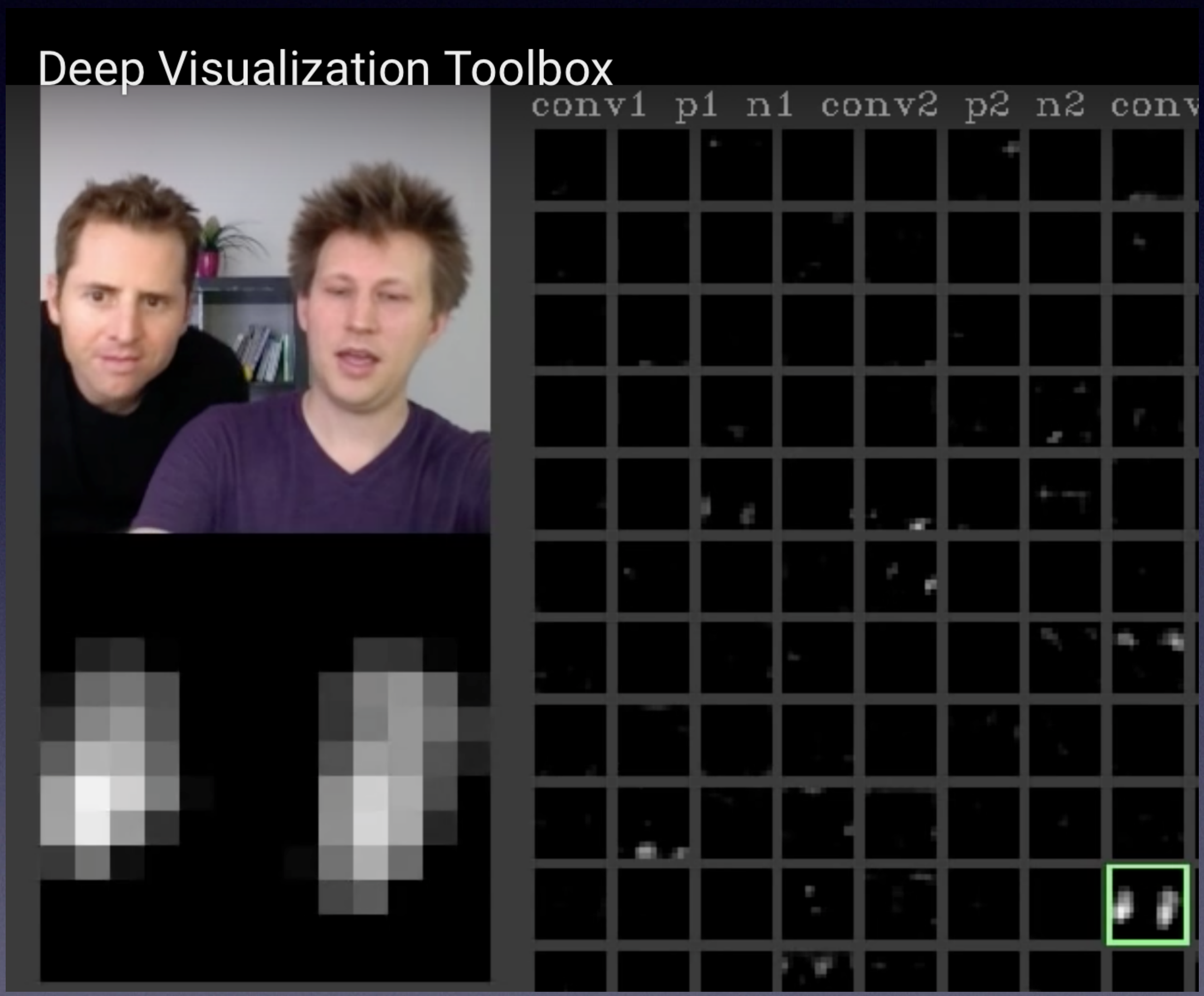
CNN for image classification

- **Goal**: provide a single label for the whole image
- **How**: transform the higher spatial resolution input (i.e. image) into a vector of image features, ultimately a 1D array of feature parameters useful for the whole image labeling, by a successful chain of convolutional and pooling operations.

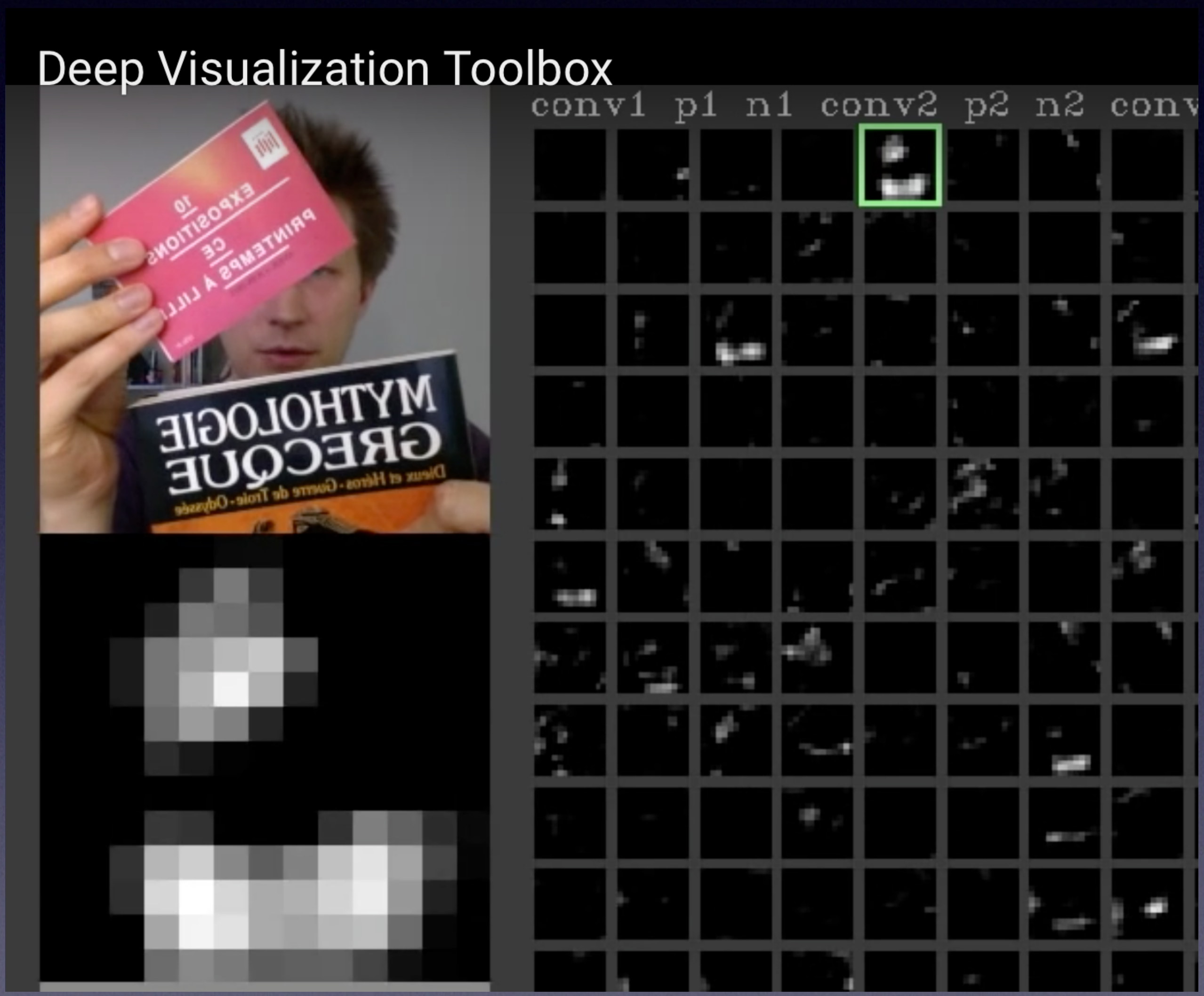
Quick Recap on CNN

Feature map visualization example

- https://www.youtube.com/watch?v=AgkfIQ4IGaM



Neuron concerning face

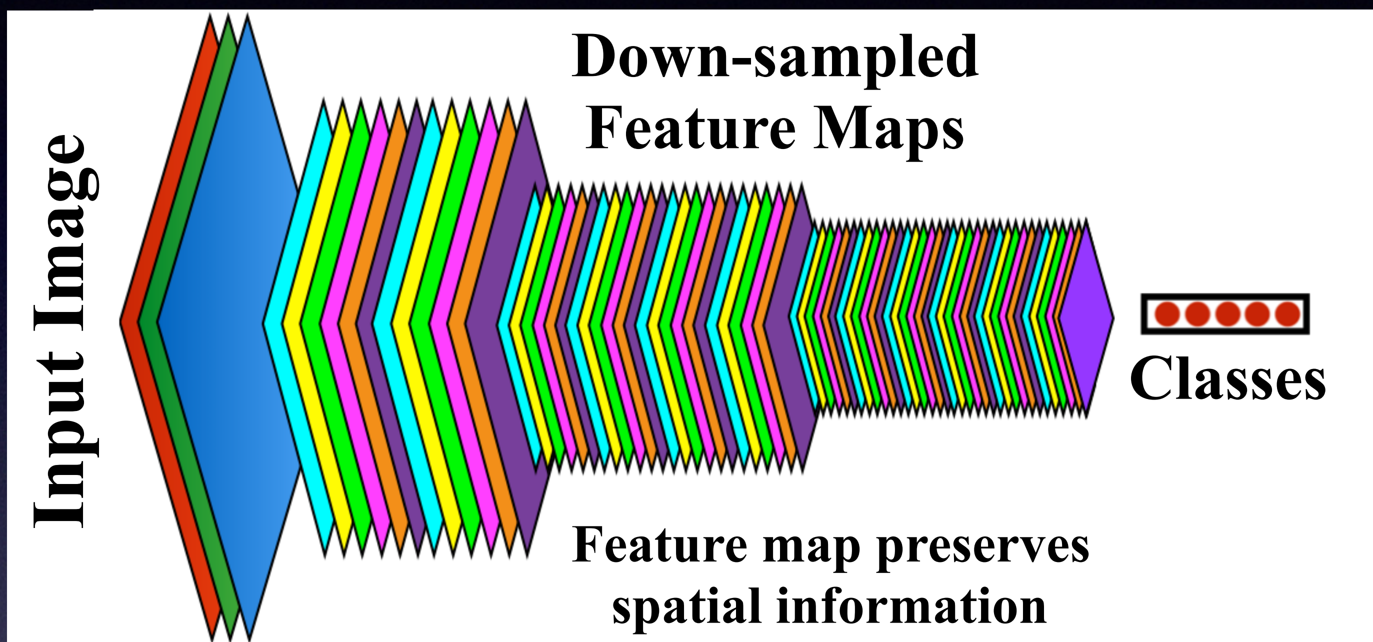


Neuron loving texts

Semantic Segmentation Network

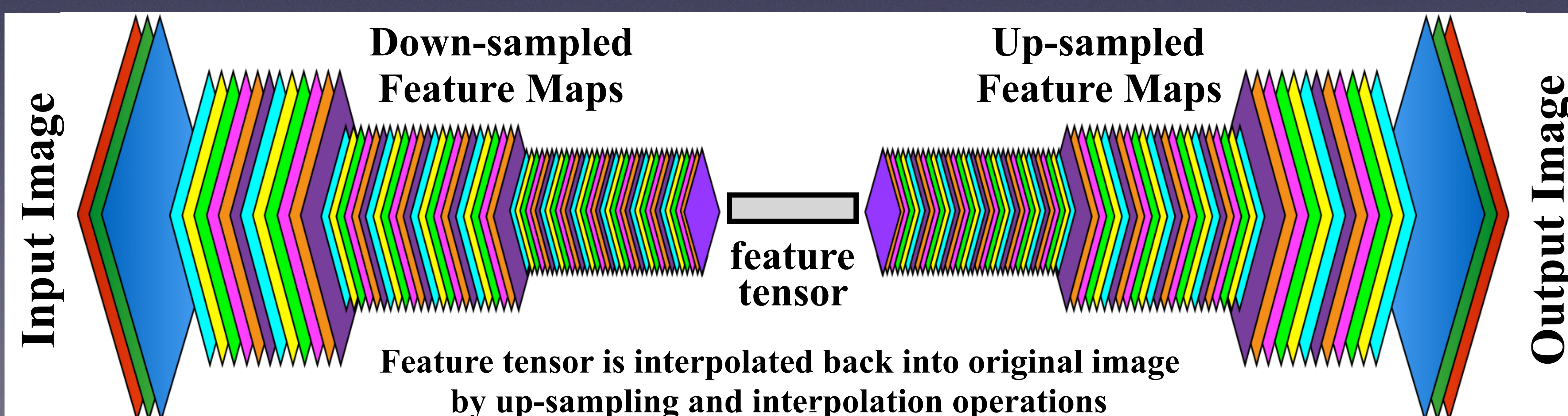
How is it different from *Image Classification*?

Example CNN for Image Classification



- Classification network reduces the whole image into final “class” 1D aray
- SSNet, after extracting class feature tensor, interpolates back into original image size

Example CNN for Semantic Segmentation

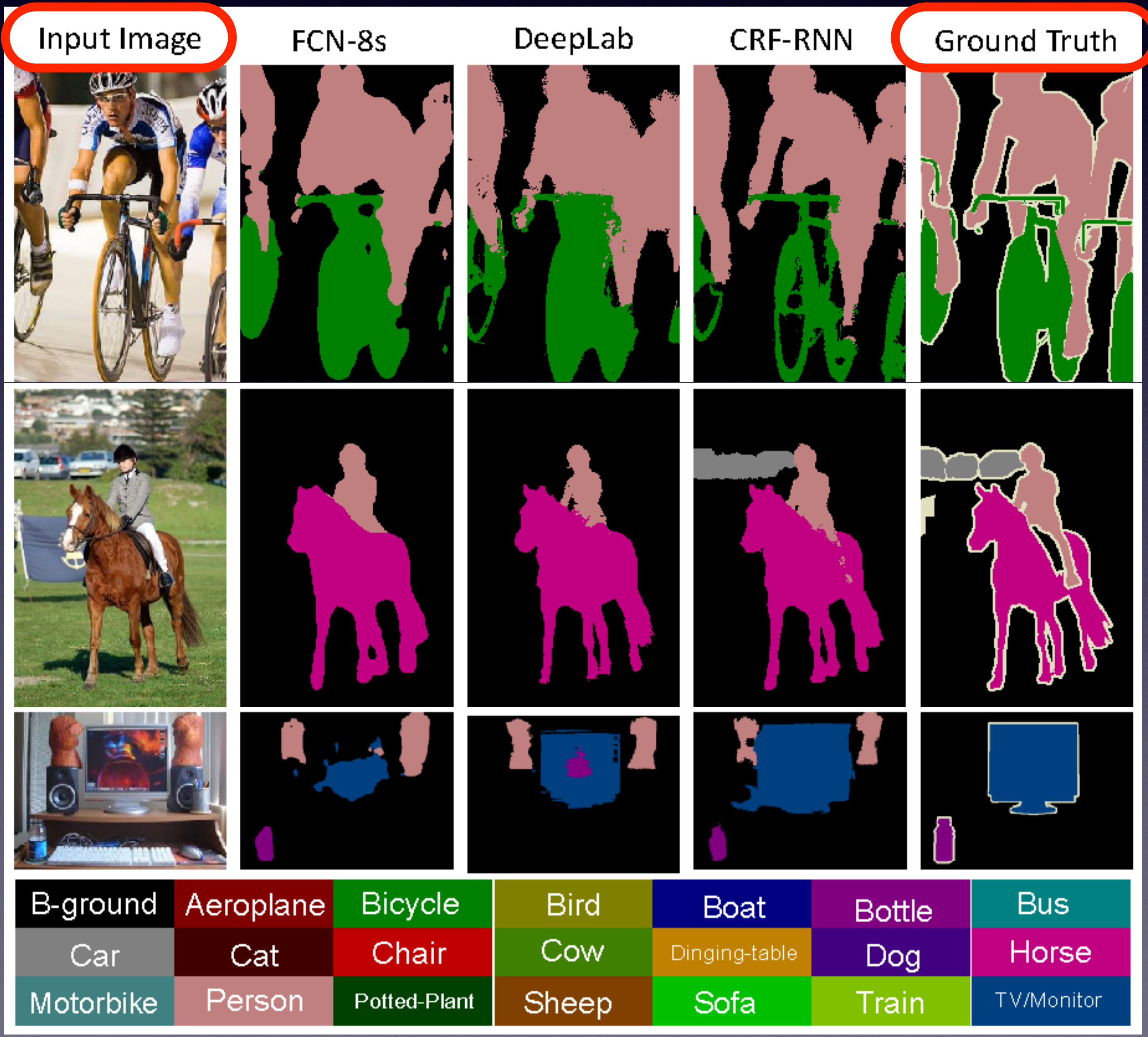


Semantic Segmentation Network

How to train SSNet?

Supervised training, like image classification

But the ***labels (and errors) are pixel-wise***



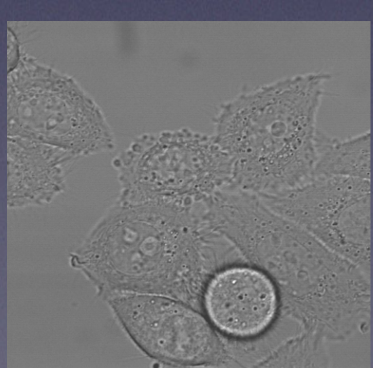
Semantic Segmentation Network

SSNet UB Analysis

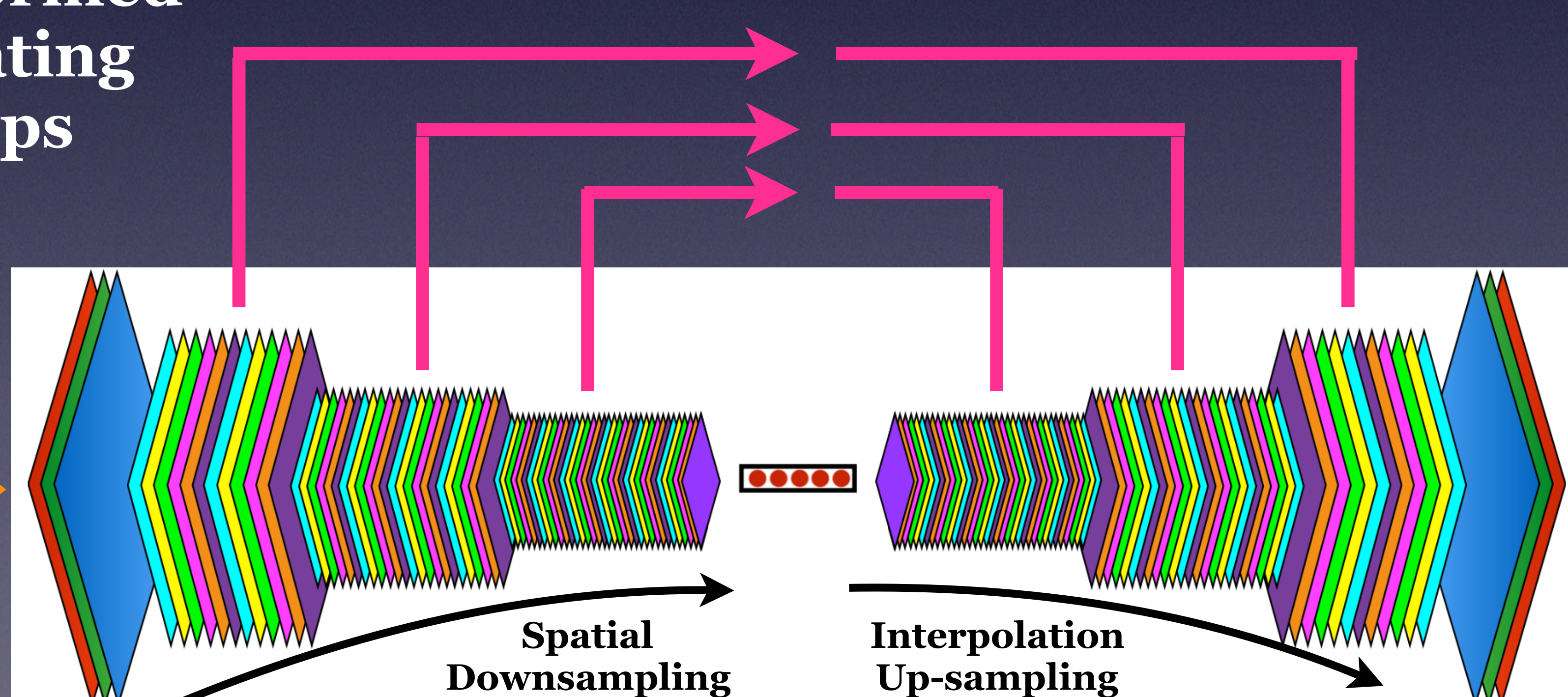
U-Net + **ResNet** module design

- Developed for bio-medical research
  - … to mask pixels of living cells (for automatized image analysis)
  - Designed for better spatial accuracy to get cells' boundary correct
- Use ResNet architecture for convolution layers

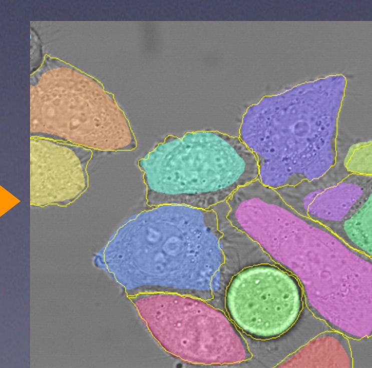
"U" shape if formed by concatenating feature maps



Biocell "Raw" Image



Segmented pixels of living cells (yellow boarder is truth label) [arXiv:1505.04597]



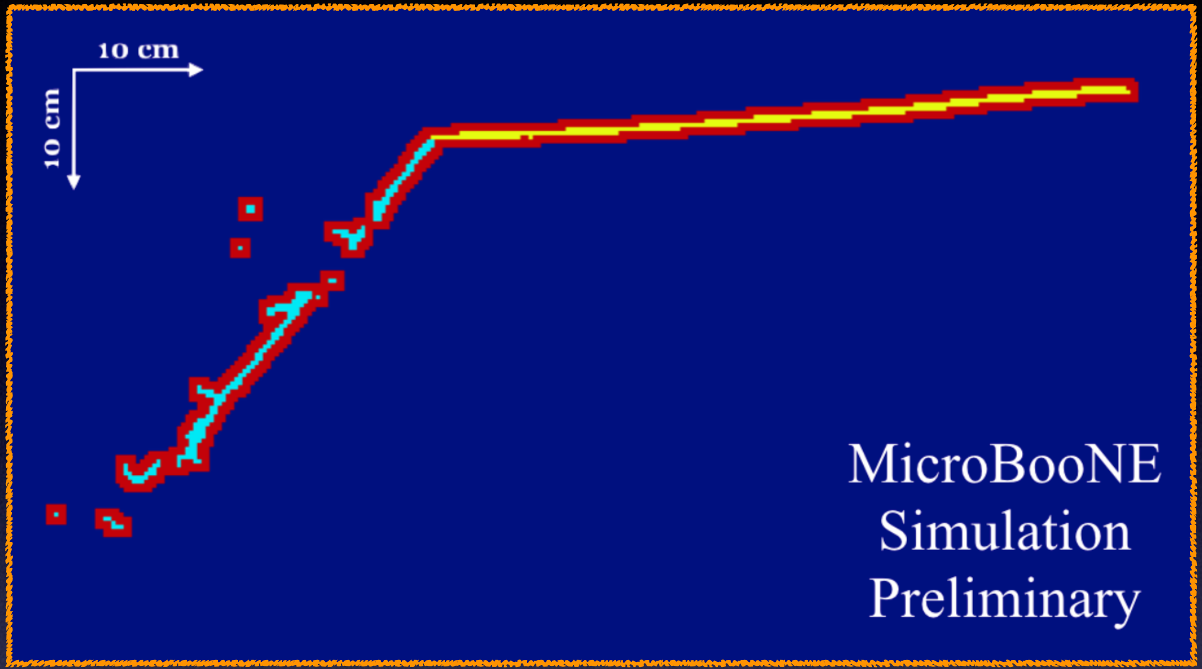
Biocell Segmented

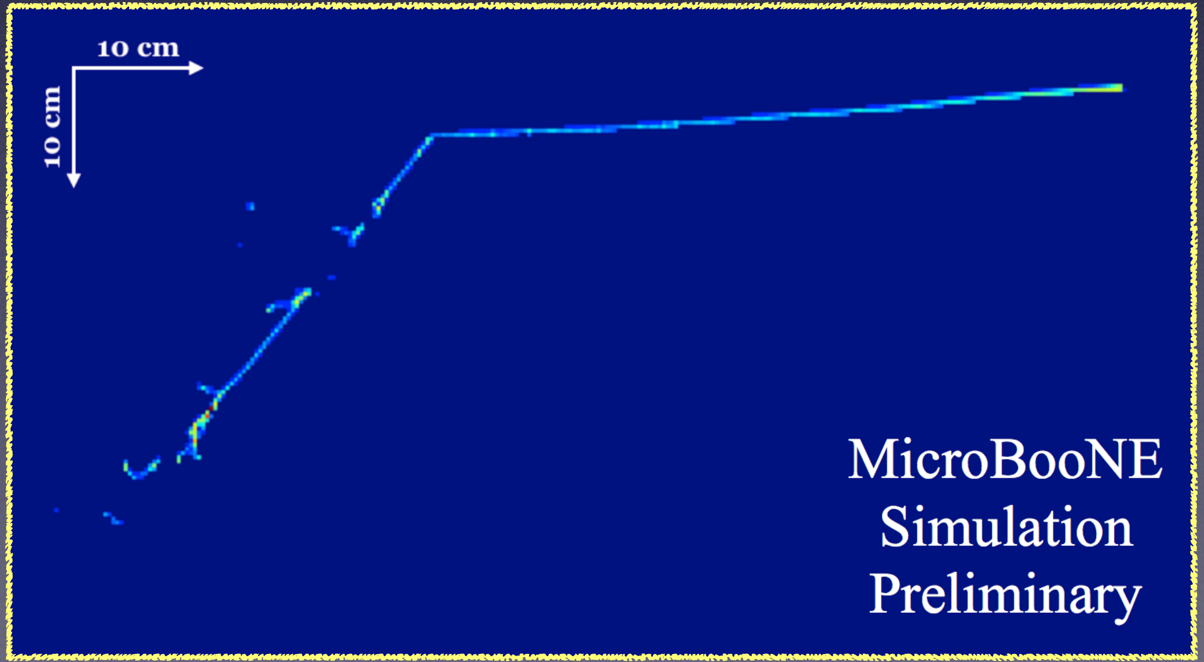
Training SSNet

“Pixel Weight” for training

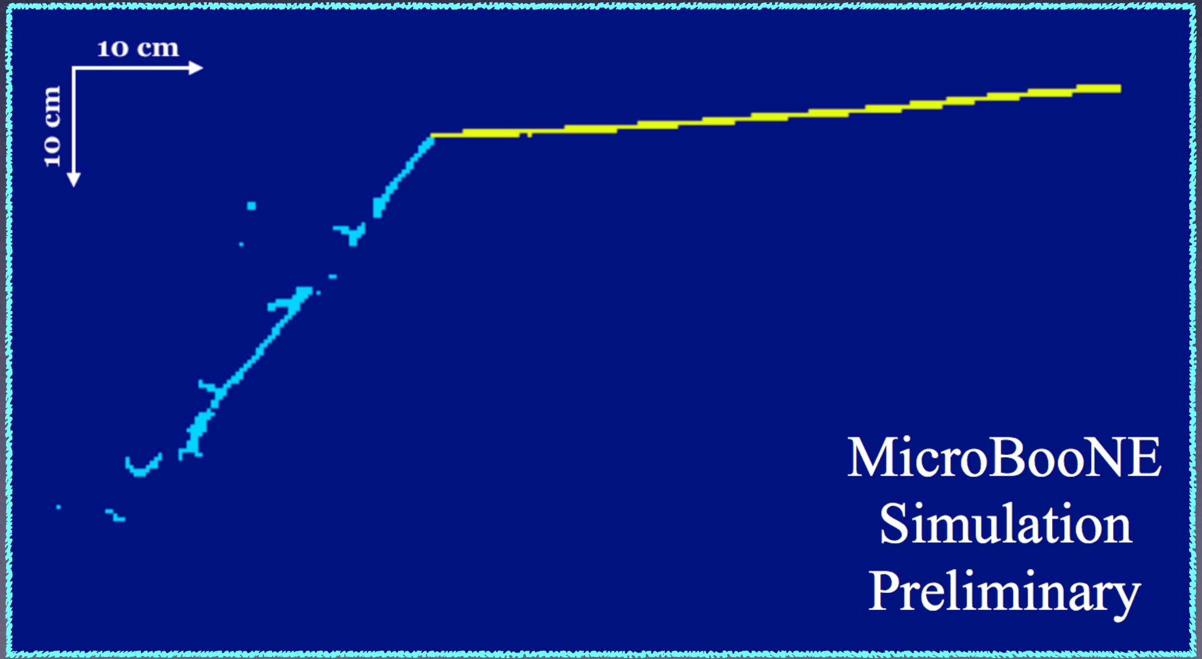
- Assign pixel-wise “weight” to penalize mistakes
- Weights inversely proportional to each type of pixel count
- Useful for LArTPC images (low information density)

“Weight” Image
(for training)



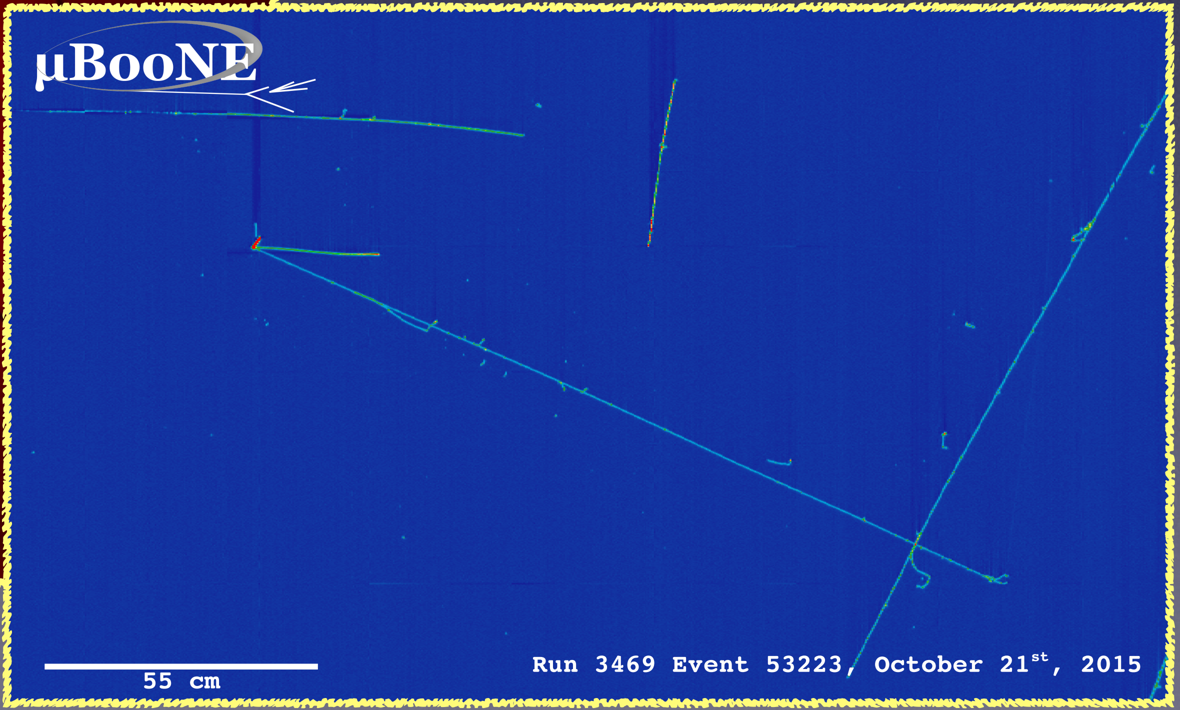


Input Image



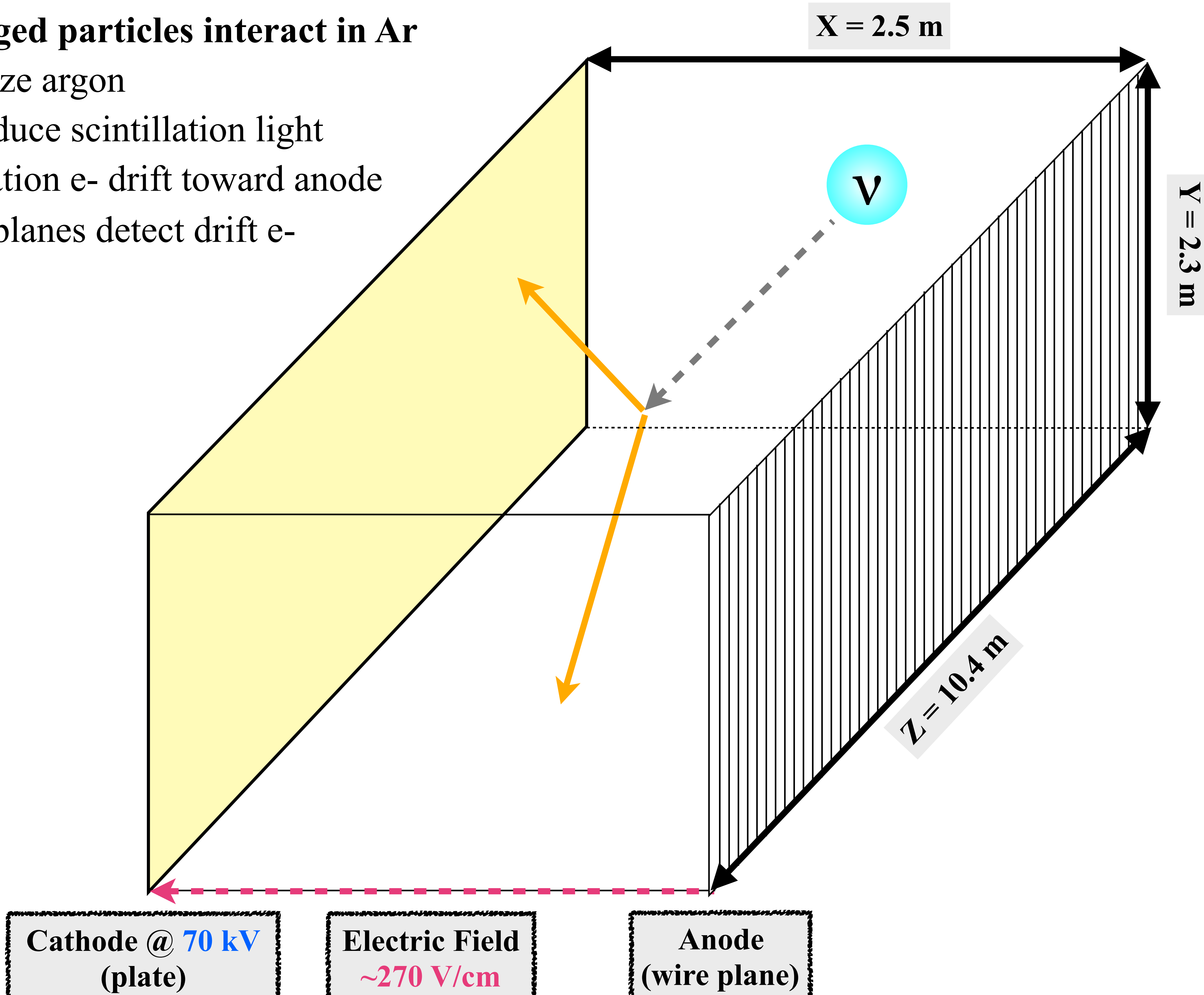
“Label” Image
(for training)

MicroBooNE LArTPC Detector Quick Guide



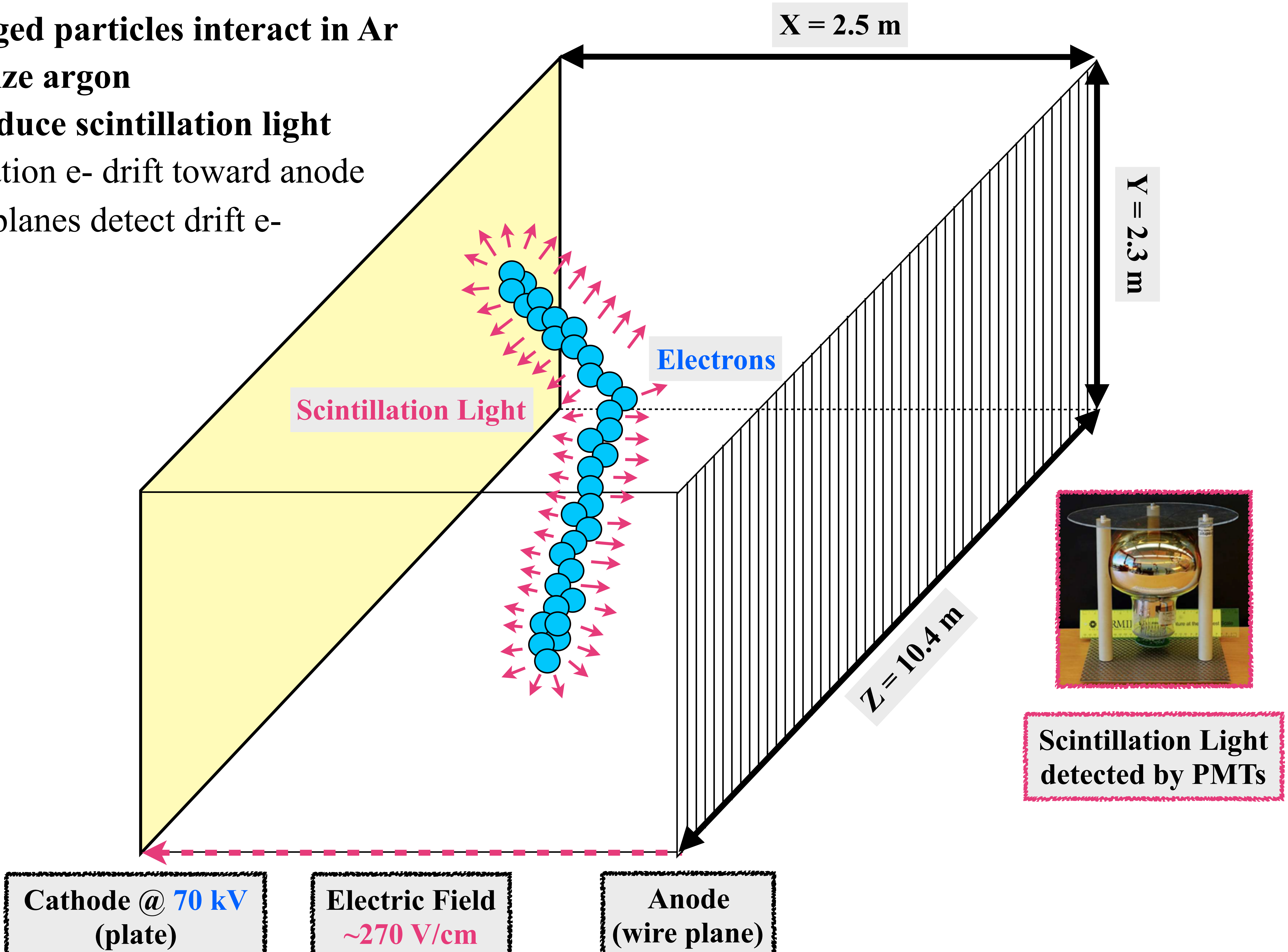
TPC Working Principle (I)

1. **Charged particles interact in Ar**
    - Ionize argon
    - Produce scintillation light
2. Ionization e- drift toward anode
3. Wire planes detect drift e-



TPC Working Principle (II)

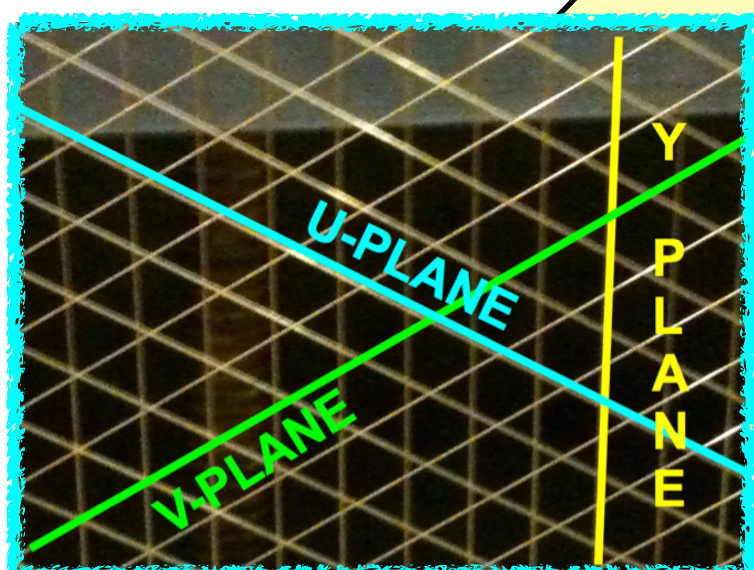
1. **Charged particles interact in Ar**
    - **Ionize argon**
    - **Produce scintillation light**
2. Ionization e- drift toward anode
3. Wire planes detect drift e-

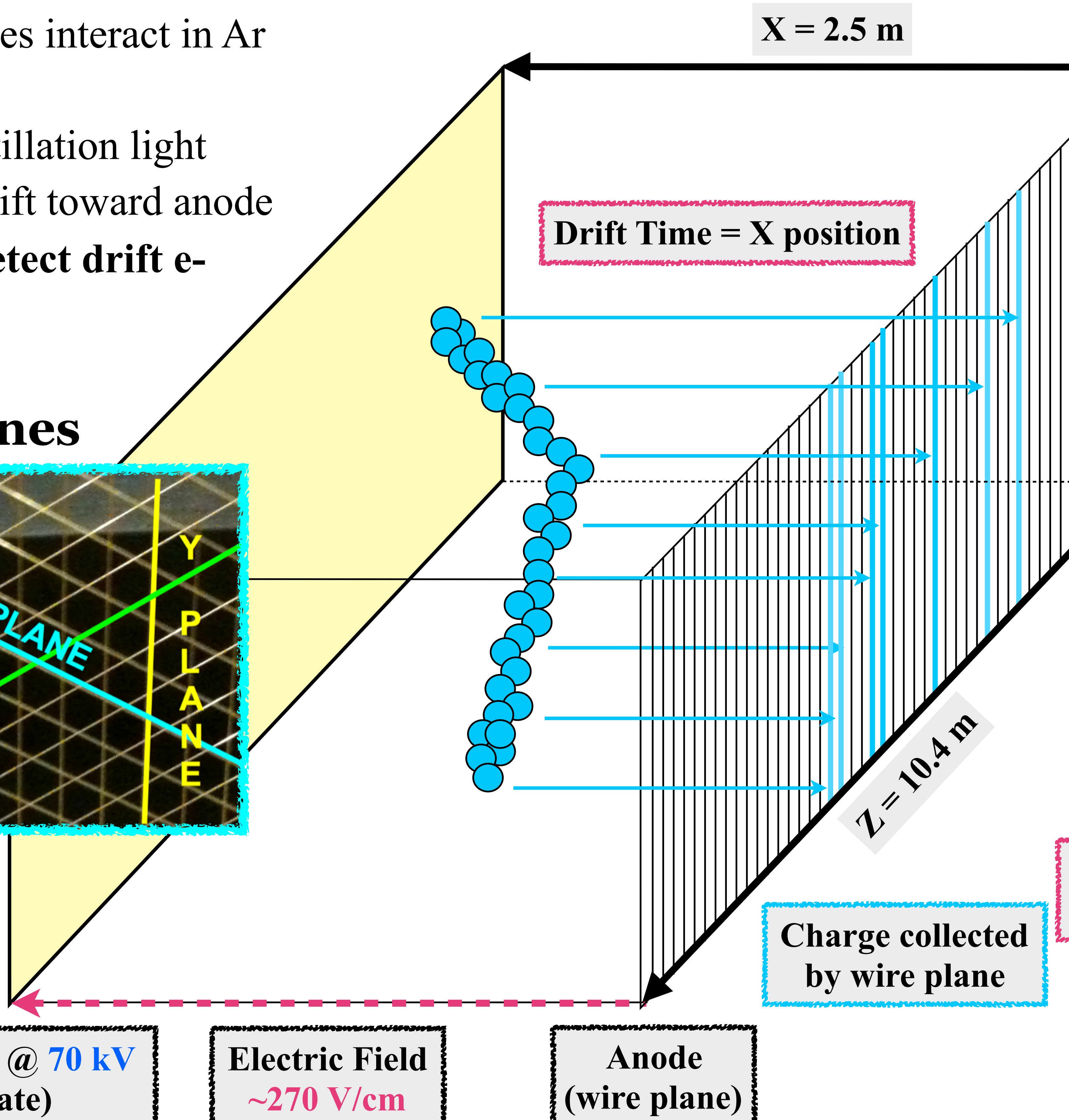


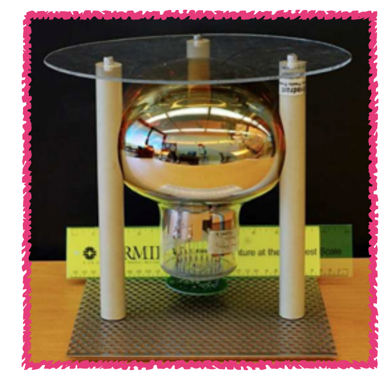
TPC Working Principle (IV)

1. Charged particles interact in Ar
    - Ionize argon
    - Produce scintillation light
2. Ionization e- drift toward anode
3. **Wire planes detect drift e-**

Three Wire Planes







Scintillation Light detected by PMTs

MicroBooNE TPC & Cryostat

Anode Wire Plane

Cathode Plate

MicroBooNE TPC & Cryostat

Anode Wir

ode Plate

MicroBooNE TPC & Cryostat